

Міністерство освіти і науки України  
Національний технічний університет України  
«Київський політехнічний інститут»

С. Р. Михайлов  
Н. В. Слободян

## **СИСТЕМИ КОНТРОЛЮ, РЕЄСТРАЦІЇ ТА ВІДОБРАЖЕННЯ ІНФОРМАЦІЇ**

**Цифрові телевізійні системи**

**Навчальний посібник**

Гриф надано Методичною радою НТУУ «КПІ»

Київ-2010

УДК

Системи контролю, реєстрації та відображення інформації.  
Цифрові телевізійні системи. Навчальний посібник для студ. спец.  
«Електронні прилади та пристрої» / Уклад.: С.Р.Михайлов, Н.В.  
Слободян. - К.: НТУУ «КПІ», 2010. - 148 с.

*Гриф надано Методичною радою НТУУ «КПІ»*

*(Протокол № 1 від 10.09.2009 р).*

Викладено основні принципи побудови систем цифрового телебачення. Розглянуто дискретизацію та квантування телевізійних сигналів та чинні стандарти щодо параметрів цих операцій. Описано методи стиснення телевізійних сигналів за стандартами JPEG, MPEG-1, MPEG-2 і MPEG-4, а також інші методи стиснення сигналів, які застосовуються в системах прикладного телебачення і відозв'язку.

Подано методи цифрового оброблення телевізійних сигналів та зображень, зокрема одновимірну та просторову цифрову фільтрацію. Розглянуто цифрові методи перетворення розгортки зображень.

Навчальний посібник призначений для студентів спеціальності «Електронні прилади та пристрої», може бути корисним студентам інших спеціальностей, аспірантам і спеціалістам у галузі телебачення.

Навчальний посібник

Михайлов Сергій Ростиславович, канд. техн. наук

Слободян Ніна Вячеславівна, канд. техн. наук

СИСТЕМИ КОНТРОЛЮ, РЕЄСТРАЦІЇ

ТА ВІДОБРАЖЕННЯ ІНФОРМАЦІЇ.

ЦИФРОВІ ТЕЛЕВІЗІЙНІ СИСТЕМИ

Відповідальний редактор Л.Д.Писаренко, докт. техн. наук, проф.

Рецензенти

М.Ю. Артеменко, докт. техн. наук, проф.

Ю.І. Танигін, канд.техн.наук, доц.

## ЗМІСТ

ВСТУП.....	4
1. ЦИФРОВІ ТЕЛЕВІЗІЙНІ СИГНАЛИ .....	6
1.1. Дискретизація сигналів і зображень .....	6
1.2. Квантування сигналів і зображень .....	15
1.3. Рекомендація ITU-R BT 601 .....	19
1.4. Інші стандарти цифрового подання телевізійних сигналів .....	34
2. СТИСНЕННЯ ТЕЛЕВІЗІЙНИХ СИГНАЛІВ .....	38
2.1. Передавання цифрових телевізійних сигналів кана- лами зв'язку. Завдання стиснення відеоінформації.....	38
2.2. Дискретне перетворення Фур'є і дискретне косинусне перетворення.....	44
2.3. Стандарт JPEG .....	52
2.4. Стандарти MPEG-1 і MPEG-2 .....	64
2.5. Стандарт кодування зображень MPEG-4 .....	85
2.6. Інші стандарти стиснення зображень .....	96
2.7. Стиснення зображень із застосуванням вейвлет-перетворення .....	99
3. ЦИФРОВЕ ОБРОБЛЕННЯ ТЕЛЕВІЗІЙНИХ СИГНАЛІВ .....	107
3.1. Цифрова фільтрація одновимірних сигналів .....	107
3.2. Просторова цифрова фільтрація .....	115
3.3. Застосування цифрової фільтрації для зменшення шумів .....	118
3.4. Перетворення розгортки .....	125
СЛОВНИК ТЕРМІНІВ І СКОРОЧЕНЬ .....	139
СПИСОК ЛІТЕРАТУРИ .....	153

## ВСТУП

Цифрове телебачення – це галузь телевізійної техніки, в якій передавання, оброблення і зберігання телевізійного сигналу здійснюються у цифровій формі.

Застосування методів та засобів цифрового телебачення – це новий ступінь розвитку телевізійної техніки, який забезпечує низку переваг порівняно з аналоговим телебаченням [1], а саме:

- підвищення перешкодостійкості телевізійних систем;
- підвищення якості зображення в телевізійних системах зі звичайним стандартом розкладення;
- створення нових телевізійних систем, що забезпечують істотне підвищення якості телевізійних зображень (телебачення високої чіткості – ТВЧ);
- створення можливості передавання стандартним телевізійним каналом із шириною смуги частот 6...8 МГц чотирьох і більше зображень звичайної чіткості або 1 – 2 зображення високої чіткості;
- забезпечення захисту телепередач і іншої інформації від несанкціонованого доступу;
- створення інтерактивних телевізійних систем.

Ці переваги зумовлені як самими принципами, притаманними цифровим телевізійним системам, так і наявністю різноманітних алгоритмів, схемних рішень та потужної технологічної бази для створення таких систем.

Цифрове оброблення телевізійних сигналів дає можливість вирішувати завдання, вирішення яких неможливе в аналогових телевізійних системах. Цифрові технології дають змогу реалізувати у телевізійних системах такі функції:

- цифрова фільтрація сигналів, зокрема для зменшення впливу

перешкод та шумів на телевізійне зображення, для вилучення відбитих сигналів тощо;

- перетворення стандартів телевізійної розгортки, наприклад для зменшення помітності мерехтінь, для реалізації режиму «кадр у кадрі» тощо;

- кодування телевізійних сигналів для зменшення швидкості їх передавання у цифровому вигляді (стиснення або компресія телевізійних сигналів);

- кодування телевізійних сигналів для зменшення впливу перешкод.

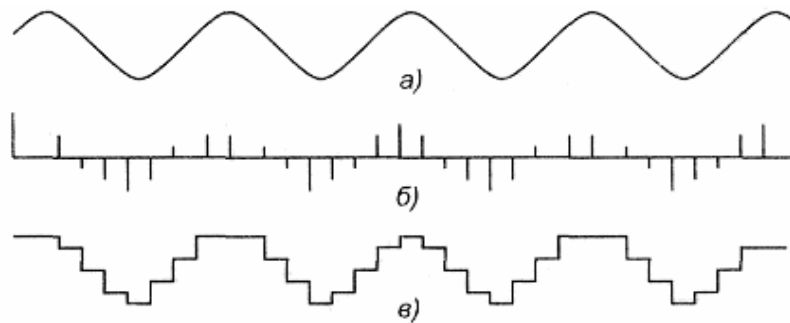
Завдяки цьому цифрові телевізійні системи найближчим часом майже повністю замінять аналогові. У той же час, у навчальній літературі ця галузь науки і техніки висвітлена недостатньо.

Навчальний посібник допоможе компенсувати недостатню кількість навчальної літератури з цієї галузі та може бути доповненням до наявної літератури з цифрових телевізійних систем. Мета посібника – ознайомити студентів із принципами побудови і роботи цифрових телевізійних систем, створених в останні роки.

# 1. ЦИФРОВІ ТЕЛЕВІЗІЙНІ СИГНАЛИ

## 1.1. Дискретизація сигналів і зображень

Дискретизація сигналу у часі – це перетворення неперервного аналогового сигналу в послідовність його значень у дискретні моменти часу. Ці значення називають відліками або вибірками [1]. В результаті дискретизації неперервного вхідного синусоїдального сигналу (рис. 1.1, а) утворюється дискретний сигнал, показаний на рис. 1.1, б. Обернене перетворення дискретного сигналу в неперервний здійснюється за допомогою операції, яку називають *інтерполяцією*, у результаті якої проміжки між відліками заповнюються за певним правилом. Найпростішу східчасту інтерполяцію зображено на рис. 1.1, в.



**Рис. 1.1.** Дискретизація та інтерполяція сигналу:

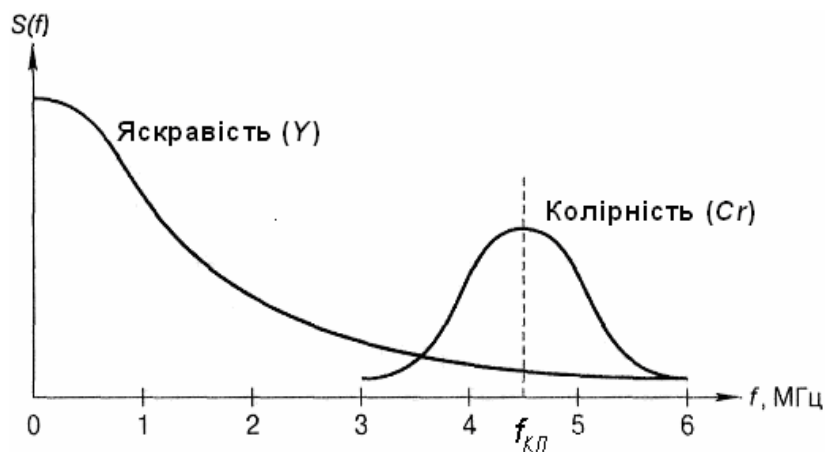
*a* – вхідний неперервний синусоїдальний сигнал; *б* – дискретний сигнал; *в* – східчаста інтерполяція дискретного сигналу

Частота дискретизації  $f_d$  має задовольняти нерівність  $f_d > 2f_v$ , де  $f_v$  – верхня гранична частота дискретизованого сигналу. Цю умову, експериментально встановлену Найквістом, була обґрунтована В. А.

Котельниковим у своїй теоремі (далі називатимемо її умовою Котельникова). В цьому прикладі вказана нерівність виконана. Тому сигнал після дискретизації та інтерполяції має порівняно незначні відмінності від вхідного сигналу, а найголовніше – має таку ж частоту, як і вхідний неперервний сигнал. Якщо під час інтерполяції дискретний сигнал пропустити через ідеальний фільтр нижніх частот (ФНЧ) з частотою зрізу, що дорівнює половині частоти дискретизації, то отриманий сигнал не матиме спотворень порівняно з вхідним сигналом.

Аналіз дискретизації та умов виникнення спотворень можливий з використанням спектрального підходу. Будь-який дійсний сигнал містить різні частотні складові. *Спектр* сигналу (або, точніше, спектральна густина) – це функція, що показує співвідношення рівнів частотних складових сигналу.

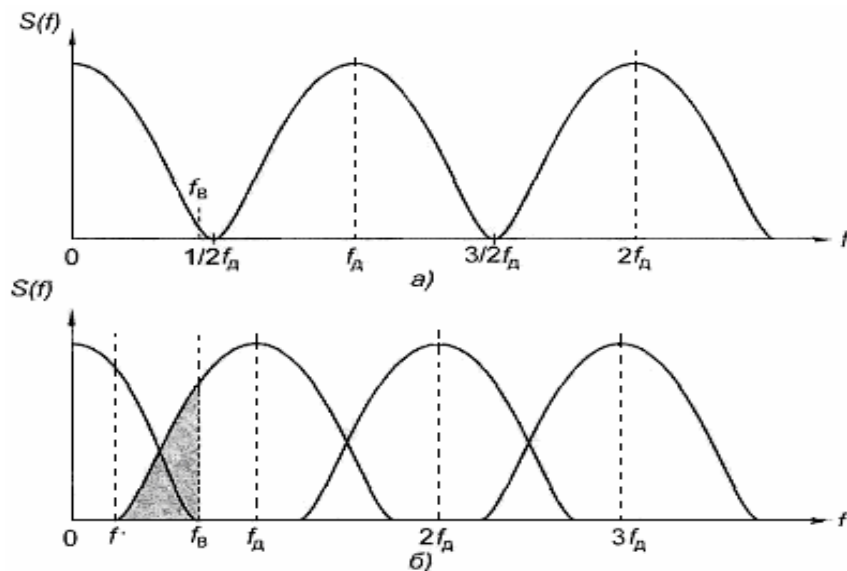
Приблизний вигляд спектра  $S(f)$  повного кольорового телевізійного сигналу (ПКТС) систем SECAM або PAL показано на рис. 1.2 [2].



**Рис. 1.2.** Спектр повного кольорового телевізійного сигналу

Спектр сигналу яскравості займає смугу частот від 50 Гц (у масштабі рисунка співпадає з 0 Гц) до 5...6 МГц. У ньому переважають низькі частоти, які відповідають великим деталям зображення. Спектр сигналу колірності розташований навколо частоти колірної піднесучої  $f_{\text{кп}}$ , і його межі перебувають на відстані приблизно 1,5 МГц від неї. Слід зазначити, що на рис. 1.2 показано *обвідні* спектрів, які насправді складаються з безлічі піків, розділених проміжками.

Дію дискретизації на спектр сигналу можна пояснити рис. 1.3.



**Рис. 1.3.** Спектри дискретизованих сигналів:

*a* – при виконанні умови Котельникова; *б* – якщо умова Котельникова не виконується

Спектр вхідного сигналу займає смугу частот від 0 до  $f_B$ . У результаті дискретизації у спектрі виникають побічні складові, які співпадають за формою із вхідним спектром і симетричні відносно



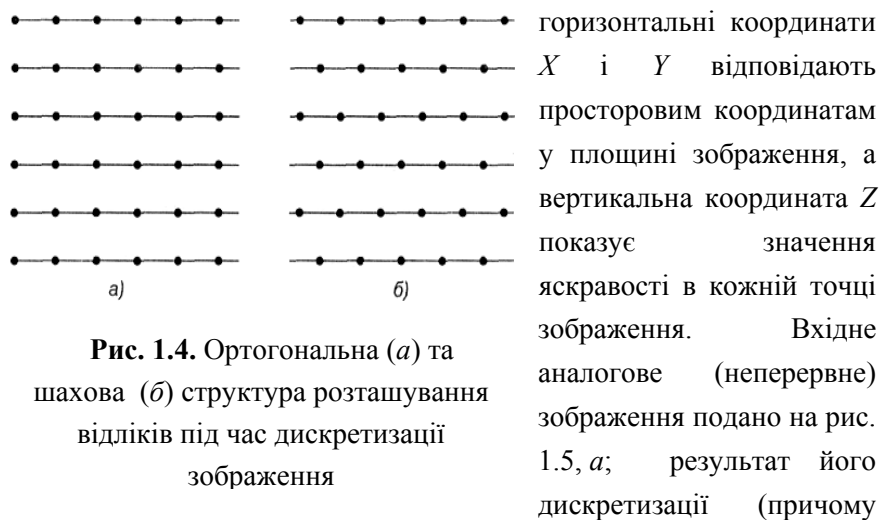
частот  $f_d, 2f_d, 3f_d, \dots$ . Якщо умова Котельникова виконується і  $f_b < f_d/2$  (рис. 1.3, а), то побічні складові спектра не перекриваються з вхідним спектром. Тому за допомогою ідеального ФНЧ, що має частоту зрізу  $f_d/2$ , можна виділити частотні складові вхідного сигналу, повністю вилучивши побічні частотні складові. Це означає, що вхідний сигнал можна відновити без спотворень.

Спектр дискретизованого сигналу у випадку, коли умова Котельникова не виконується, тобто  $f_d < 2 f_b$ , показано на рис. 1.3, б. При цьому спектр вхідного сигналу і спектр хоча б однієї з побічних складових, що виникли під час дискретизації, перекриваються. Це явище не можна усунути жодним фільтром, тому спотворення, причинені дискретизацією, у цьому випадку необоротні. Якщо, наприклад, намагатися виділити вхідний сигнал за допомогою ідеального ФНЧ з частотою зрізу  $f_b$  або  $f_d/2$ , то на вихід цього ФНЧ, крім вхідного сигналу, частково пройде побічна складова. Якщо ж використовувати ідеальний ФНЧ з частотою зрізу  $f' = f_d - f_b$ , то побічна складова не пройде, але і високочастотна частина спектра вхідного сигналу буде вилучена, що також призведе до спотворення сигналу.

Щоб уникнути необоротних спотворень під час дискретизації, потрібно обирати достатньо велику частоту дискретизації та обмежувати верхню граничну частоту дискретизованого сигналу так, щоб умова Котельникова виконувалася.

Далі розглянемо дискретизацію нерухомих зображень [3, 4], за якої відліки утворюють певну структуру у площині зображення. Найбільш поширену ортогональну (прямокутну) структура зображено на рис. 1.4, а, а шахову структуру – на рис. 1.4, б.

Процес дискретизації зображення та його подальшого відтворення можна ілюструвати також тривимірними діаграмами (рис. 1.5), де



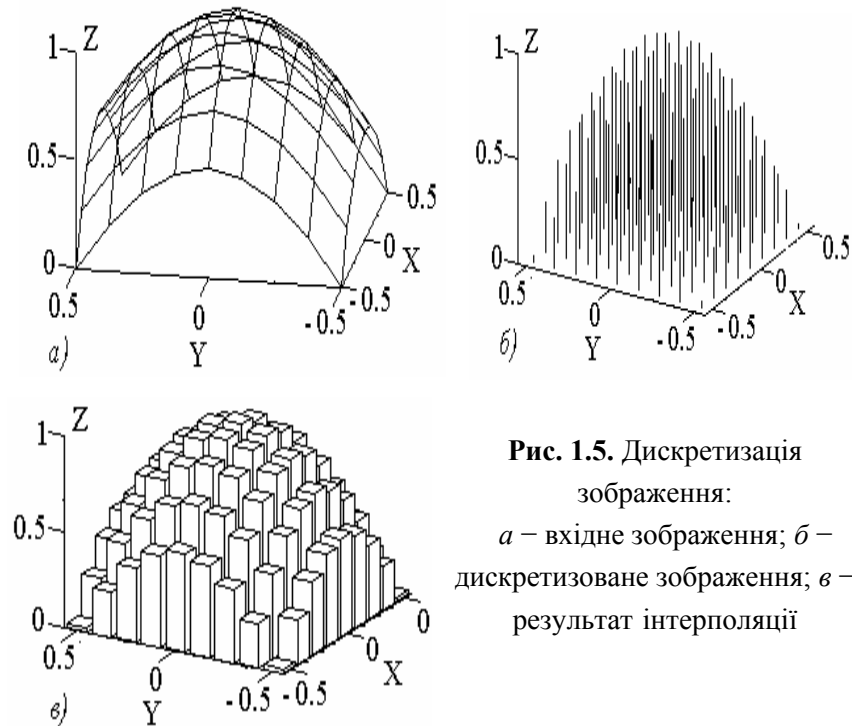
**Рис. 1.4.** Ортогональна (а) та шахова (б) структура розташування відліків під час дискретизації зображення

відліки розташовані у вузлах ортогональної ґратки) – на рис. 1.5, б; двовимірна функція яскравості, що утворюється під час відтворення дискретизованого зображення на екрані з використанням двовимірної східчастої інтерполяції, – на рис. 1.5, в. Відтворене зображення складається з прямокутних елементів, які називають пікселами (pixel – picture element). Яскравість і колір у межах будь-якого пікселя майже постійні. Розміри пікселів мають бути настільки малими, щоб глядач за нормальних умов спостереження сприймав відтворене зображення як неперервне.

Як і у разі одновимірного сигналу, можна розглянути дискретизацію із спектральної точки зору. Для цього спочатку введемо поняття «просторова частота» і «просторовий спектр» для неперервного зображення.

За аналогією з періодом одновимірного сигналу, що є функцією часу, назовемо просторовим періодом  $T_x$  за координатою  $X$  просторовий інтервал, через який значення двовимірного сигналу  $b(x, y)$  повторюються. Тоді просторовою частотою цього сигналу за

координатою  $X$  буде величина, обернена просторовому періоду  $f_x = 1/T_x$ . Аналогічно вводяться просторовий період і просторова частота за координатою  $Y$ .

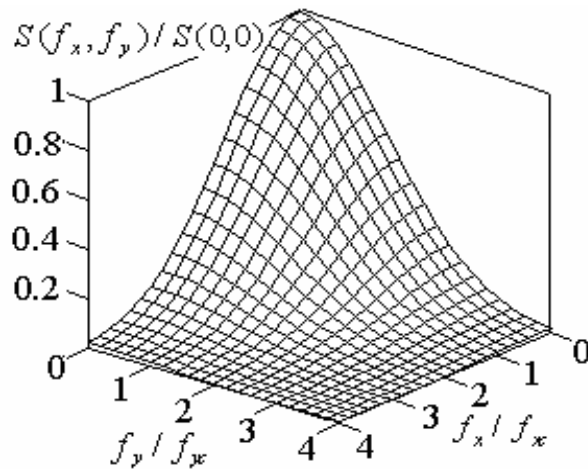


**Рис. 1.5.** Дискретизація зображення:  
 $a$  – вхідне зображення;  $б$  – дискретизоване зображення;  $в$  – результат інтерполяції

Низькі просторові частоти відповідають повільним змінам яскравості зображення за відповідними просторовими координатами, а високі просторові частоти – швидким змінам яскравості, тобто дрібним деталям зображення і різким перепадам яскравості на межах об'єктів.

Наближений вигляд просторового спектра  $S(f_x, f_y)$  реального зображення показано на рис. 1.6, в якому переважають низькі просторові частоти. На горизонтальних осях відкладено значення

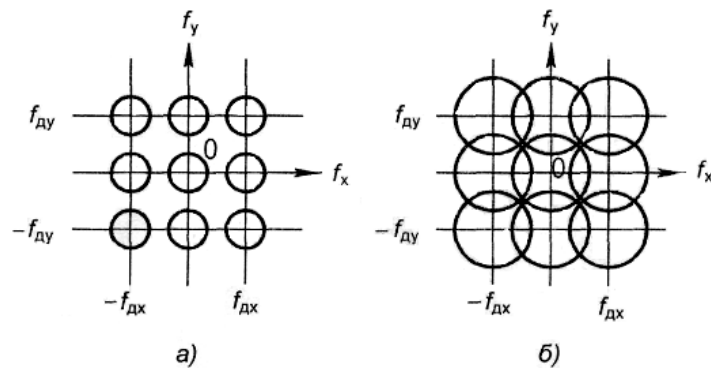
відношень просторових частот до значень просторових частот за відповідними координатами, для яких модуль спектра зменшується в  $e = 2,718$  разу відносно постійної складової  $S(0, 0)$ . По вертикальній осі графік поділений на цю постійну складову.



**Рис. 1.6.** Наближений вигляд просторового спектра реальних зображень

Перетворення просторового спектра зображення у разі двовимірної дискретизації подано на рис. 1.7. Просторовий спектр вхідного зображення передбачається обмеженням у площині просторових частот, тобто ззовні деякої замкненої кривої всі частотні складові можна вважати рівними нулю. Як і у разі дискретизації одновимірних сигналів, під час двовимірної дискретизації з'являються побічні складові спектра, зміщені відносно вхідного спектра по горизонталі і по вертикалі на значення просторових частот дискретизації  $f_{dx}$  і  $f_{dy}$  відповідно.

Випадок, коли в результаті дискретизації побічні складові не



**Рис. 1.7.** Просторові спектри дискретизованих зображень у випадках виконання (а) та порушення (б) умов теореми Котельникова для двовимірних сигналів

перетинаються з вхідним спектром, показано на рис. 1.7, а. Це досягається за достатньо великих просторових частотах дискретизації за обома координатами. За таких умов можливе відновлення вхідного зображення за дискретизованим із застосуванням просторового ФНЧ, що виділяє спектр вхідного зображення зі спектра дискретизованого зображення. Це є узагальненням теореми Котельникова на двовимірні сигнали.

Випадок, коли побічні складові спектра дискретизованого зображення перекриваються зі спектром вхідного зображення, що є наслідком недостатньо великих частот дискретизації за просторовими координатами, показано на рис. 1.7, б. Відновлення вхідного зображення за дискретизованим без спотворень за таких умов неможливе.

Конкретний вид спотворень у зображенні залежить від форми об'єктів, з яких воно складається, та особливостей його просторового спектра. За наявності у вхідному зображенні добре виражених періодичних складових спотворення можуть виявитися у

вигляді нових періодичних складових (муарові спотворення).

Щоб спотворень не виникало, потрібно обирати достатньо великі просторові частоти дискретизації за обома координатами. Проте в телебаченні ці просторові частоти фактично визначені параметрами, встановленими стандартом розкладення, тобто кількістю рядків і елементів у кожному рядку.

Для узгодження просторового спектра зображення із вказаними параметрами у багатьох випадках доводиться обмежувати верхні граничні просторові частоти зображення перед дискретизацією. Цю операцію виконують за допомогою спеціальних оптичних розсіювальних елементів, розташованих перед матрицею на основі приладів із зарядовим зв'язком (ПЗЗ-матрицею), або шляхом незначного розфокусування об'єктива у телевізійній камері.

У телебаченні передають зображення рухомих об'єктів. Звичайне, тобто аналогове телевізійне зображення, вже є дискретним за часом і за вертикальною координатою, оскільки складається з послідовності кадрів і кожен кадр розкладений на рядки. Ці дві дискретизації виконуються безпосередньо в перетворювачі світло–сигнал (в ПЗЗ-матриці або в передавальній трубці). Щоб зробити зображення дискретним за горизонтальною координатою, виконується одновимірна дискретизація телевізійного сигналу в часі.

Якщо задавати частоту дискретизації телевізійного сигналу за його верхньою граничною частотою (звичайною, а не просторовою), відповідно до умови Котельникова, то задовольнятиметься і двовимірною умовою, коли немає спотворень зображення за рахунок перекриття просторових спектрів під час дискретизації. Таким чином, частота дискретизації телевізійного сигналу, що має за мовним стандартом розкладення зображення верхню граничну частоту близько 6 МГц, повинна бути не меншою за 12 МГц.

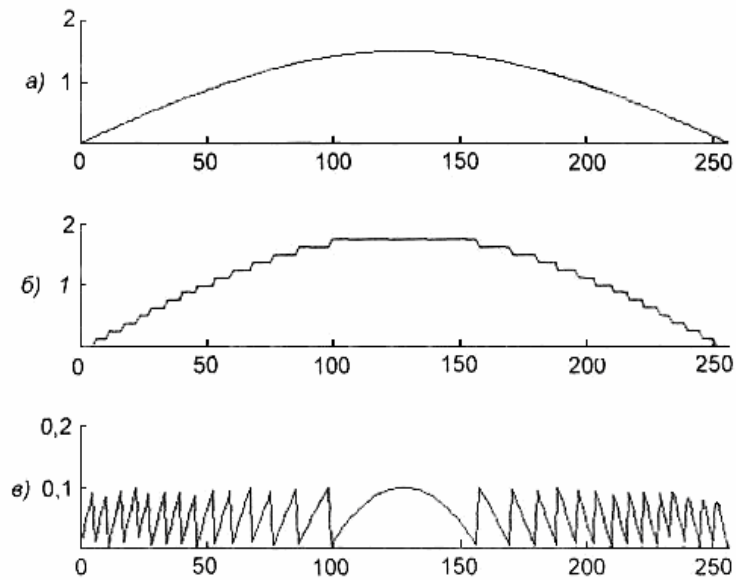
## 1.2. Квантування сигналів і зображень

Квантування сигналу за рівнем полягає у визначенні для кожного відліку сигналу найближчого до нього рівня квантування серед рівнів квантування, які використовують, тобто заміни значень відліків сигналу найближчими рівнями квантування. Кількість рівнів квантування  $N_{\text{кв}}$  пов'язана з числом двійкових розрядів ( $b$ ) аналого-цифрового перетворювача (АЦП) співвідношенням  $N_{\text{кв}} = 2^b$ .

Розглянемо приклад квантування сигналу. Дискретний сигнал, що містить 256 відліків (окремі відліки в цьому масштабі не помітні), показано на рис. 1.8, *а*; значення рівня сигналу вимірюються у вольтах і змінюються неперервно, тобто квантування немає. Той же сигнал після квантування на 16 рівнів з числом двійкових розрядів  $b = 4$ , показано на рис. 1.8, *б*. Розмір кроку квантування  $h = 0,1$ . Масштаб шкали по вертикальній осі координат такий, як на попередньому рисунку.

Значення похибки  $\Delta U$ , що вноситься в сигнал операцією квантування, показано на рис. 1.8, *в*. При цьому передбачається, що квантування здійснюється за рахунок заміни значення сигналу найближчим знизу рівнем квантування. Значення похибки квантування  $\Delta U$  змінюється в межах від 0 до  $h$ . Якщо ж замінювати сигнал найближчим зверху або знизу рівнем квантування, то значення  $\Delta U$  змінюватиметься в діапазоні  $-h/2 \dots h/2$ .

Похибка квантування є випадковою величиною, тому її часто називають шумом квантування. У випадку, коли всі значення сигналу рівноймовірні, розподіл імовірності значення шуму квантування також рівномірний. Середнє квадратичне значення шуму квантування у разі квантування замінюванням найближчим рівнем, постійного кроку квантування і рівноймовірних значень сигналу дорівнює  $\sigma_{\text{кв}} = h/\sqrt{12}$ .



**Рис. 1.8.** Сигнал до квантування (а), сигнал після квантування (б), похибка квантування (в)

Дисперсія шуму квантування, яка характеризує його потужність, може бути виражена в логарифмічному масштабі (у децибелах) відносно пікового рівня потужності вхідного сигналу, дБ,

$$\sigma_{\text{кв}}^2 = 10 \lg \left( \frac{h^2}{12} \right) = 10 \lg \left( \frac{2^{-2b}}{12} \right) \approx -(6,02b + 10,8),$$

де  $b$  – кількість двійкових розрядів без урахування знакового розряду; максимальний рівень сигналу взято рівним одиниці.

Якщо повна кількість розрядів АЦП дорівнює 8, а з них один біт визначає полярність сигналу, то  $b = 7$  і  $\sigma_{\text{кв}}^2 = -53$  дБ, тобто середня потужність шуму квантування майже у  $2 \cdot 10^5$  разів менше за максимальну потужність корисного сигналу. За умови, що АЦП 16-



розрядний, АЦП маємо  $b = 15$ , а  $\sigma_{\text{кв}}^2 = -101$  дБ. Слід пам'ятати, що ці значення одержані для максимального рівня сигналу. У разі зменшення рівня сигналу відношення сигнал/шум зменшуватиметься.

На зображенні шум квантування може виявлятися по різному залежно від характеру змін яскравості або кольору на цій ділянці зображення. На ділянках, що складаються з дрібних деталей, квантування призведе до випадкових змін їх яскравості або кольору. На ділянках зображення з плавною зміною рівня відеосигналу квантування може призвести до появи хибних контурів на тих лініях, на яких рівень відеосигналу переходить від одного рівня квантування на інший.

Рівномірне квантування телевізійного сигналу не є найкращим. Це обумовлено, у першу чергу, властивостями зорового аналізатора людини. Відповідно до закону Вебера–Фехнера, порогове перевищення яскравості об'єкта над фоном  $\Delta B_{\text{пор}}$ , за яким об'єкт розрізняється спостерігачем, дорівнює 1,5...2,0 % від значення яскравості фону  $B_0$  [5]. Таким чином, зі збільшенням яскравості фону збільшується і поріг  $\Delta B_{\text{пор}}$ . Звідси випливає, що у області значень телевізійного сигналу, близьких до рівня чорного, крок квантування має бути меншим, ніж в області, близькій до рівня білого.

Проте технічна реалізація нерівномірного квантування значно складніша, ніж рівномірного. Замість змінного кроку квантування, зазвичай виконують попереднє нелінійне перетворення відеосигналу – гамма-корекцію. При цьому вирішують одночасно два завдання: по-перше, коректується нелінійність передавальної характеристики

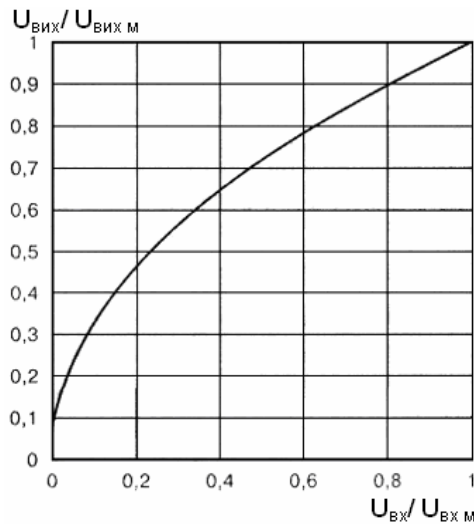
кінескопа і забезпечується оптимальна форма передавальної характеристики всього тракту телевізійної системи «від світла до світла»; по-друге, зменшується вплив похибок квантування за малих рівнів яскравості зображення.

Передавальну характеристику гамма-коректора описують співвідношенням

$$\frac{U_{\text{вих}}}{U_{\text{вих м}}} = \left( \frac{U_{\text{вх}}}{U_{\text{вх м}}} \right)^{\gamma},$$

де  $U_{\text{вх}}$  і  $U_{\text{вих}}$  – напруга сигналів на вході і на виході гамма-коректора відповідно;  $U_{\text{вх м}}$  і  $U_{\text{вих м}}$  – максимальні значення діапазонів напруг сигналів на вході і на виході гамма-коректора відповідно;  $\gamma$  – показник гамма-корекції,  $\gamma = 0,42 \dots 0,48$ . Графік передавальної характеристики гамма-коректора для  $\gamma = 0,45$

зображено на рис. 1.9.



**Рис. 1.9.** Передавальна функція гамма-коректора

У системах цифрового телебачення переважно застосовують рівномірне квантування сигналів, які пройшли гамма-корекцію. Кількість двійкових розрядів АЦП ( $b$ ) при цьому складає 8, що зумовлює кількість рівнів квантування  $N_{\text{кв}} = 256$ . За цих умов шум квантування на зображенні майже непомітний. В останні роки

все частіше застосовують квантування, де  $b = 10$ , що дає змогу поліпшити якість зображення.

На виході АЦП номер рівня квантування подається у вигляді двійкового числа, тобто виконується кодування. Якщо кількість розрядів АЦП дорівнює 8, то значення цифрового сигналу подаються числами від 00000000 до 11111111.

### **1.3. Рекомендація ITU-R BT 601**

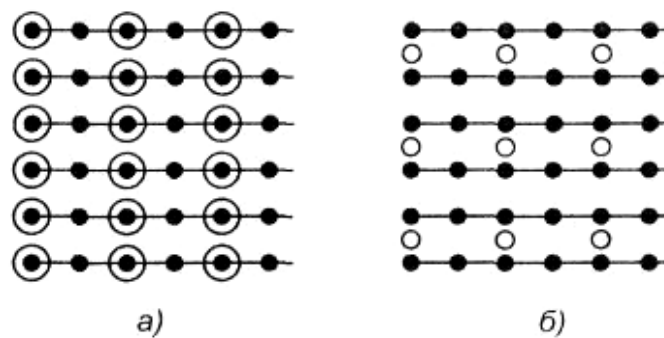
Важливу роль у розвитку цифрового телебачення відіграють міжнародні стандарти. Головним органом, що розробляє і затверджує стандарти в різних галузях людської діяльності, є Міжнародна організація по стандартизації – International Standard Organization (ISO), яка об'єднує національні комітети зі стандартизації більше 100 країн світу. ISO створює комітети, підкомітети і робочі групи, які розробляють проекти стандартів, що затверджуються потім ISO. При цьому стандарти у галузі інформаційних технологій розробляє та затверджує створений ISO і Міжнародною Електротехнічною Комісією (International Electrotechnical Commission - IEC) спільний комітет ISO/IEC JTS1, назву якого зазначено у виданих ним стандартах.

Ще один міжнародний орган, що займається стандартизацією у галузі телекомунікацій, це Міжнародний союз електрозв'язку (International Telecommunication Union – ITU). Стандартизацією в телебаченні займається сектор ITU-R (R – radio). До 1993 року ці функції виконував Міжнародний консультативний комітет по радіо (МККР або CCIR). Документи, розроблені ITU, називають Рекомендаціями, але фактично вони є стандартами.

Рекомендація ITU-R BT 601 (або Рекомендація 601) визначає єдиний міжнародний стандарт цифрового кодування телевізійного

сигналу. Цей стандарт застосовують у сучасних повністю цифрових телевізійних системах під час цифрового подання телевізійних сигналів звичайної чіткості. У цьому стандарті передбачено роздільне цифрове кодування сигналу яскравості і двох кольорорізницевих сигналів. Такий варіант кодування називають компонентним [6].

**Дискретизація.** Встановлено одне значення частоти дискретизації сигналу яскравості, яке дорівнює 13,5 МГц для обох стандартів розгортки - 625 рядків, 25 Гц і 525 рядків, 30 Гц. Кожен кольорорізницевий сигнал дискретизується з частотою 6,75 МГц. Такий формат дискретизації телевізійних сигналів позначають як 4:2:2. Це означає, що частота дискретизації кольорорізницевих сигналів у 2 рази менша, ніж частота дискретизації сигналу яскравості, причому в кожному рядку наявні обидва кольорорізницеві сигнали. Взаємне розміщення відліків яскравості і кольорорізницевих сигналів показано на рис. 1.10, а.



**Рис. 1.10.** Розміщення відліків яскравості і кольорорізницевих сигналів: *а* – формат дискретизації 4:2:2; *б* – формат дискретизації 4:2:0; • – відліки сигналів яскравості; о – відліки кольорорізницевих сигналів

Всього відліків яскравості в рядку 864, а кількість відліків кожного кольорорізнцевого сигналу – 432. За час активної ділянки рядка формується 720 відліків сигналу яскравості і 360 відліків кожного кольорорізнцевого сигналу. Ці кількості відліків є проміжними між значеннями, визначеними у вказаних стандартах розгортки для отримання квадратної форми пікселів. Отже, під час розробки Рекомендації 601 було досягнуто компромісу: кількість активних рядків в кадрі для стандарту 625 рядків становить 576, а повна кількість елементів зображення, яке передається в кожному кадрі, для сигналу яскравості становить 414720.

Передбачені також інші варіанти дискретизації телевізійних сигналів. Наприклад, для формату 4:2:0 кожен кольорорізнцевий сигнал має частоту дискретизації у 2 рази нижчу за частоту дискретизації сигналу яскравості, і кількість рядків, що передаються, зменшується удвічі. Розміщення відліків яскравості і кольорорізнцевих сигналів у площині зображення для цього варіанта показано на рис. 1.10, б. Матриця відліків для кожного з кольорорізнцевих сигналів містить  $360 \times 288$  елементів.

У форматі 4:1:1 кількість рядків для кожного кольорорізнцевого сигналу така ж, як для сигналу яскравості, але частоти дискретизації кольорорізнцевих сигналів у 4 рази менші від частот дискретизації сигналу яскравості, тобто дорівнюють 3,375 МГц. Кількість елементів кожного кольорорізнцевого сигналу в кадрі в цьому випадку така ж, як для формату 4:2:0.

Нарешті, у форматі 4:4:4 обидва кольорорізнцевих сигнали мають стільки ж рядків і дискретизуються з тією ж частотою, що і сигнал яскравості.

Під час цифрового оброблення телевізійних сигналів у відеоапаратурі (фільтри для зменшення шумів, перетворювачі розгортки) зазвичай використовують формати 4:1:1 і 4:2:2.

**Квантування.** У Рекомендації 601 для всіх сигналів передбачена кількість розрядів квантування  $b = 8$ , що визначає кількість квантування  $N_{\text{кв}} = 256$ . При цьому рівню чорного сигналу яскравості відповідає 16-й рівень квантування, а номінальному рівню білого – 235-й рівень квантування. 16 рівнів квантування знизу і 20 рівнів квантування зверху утворюють резервні зони на випадок виходу значень аналогового сигналу яскравості за межі номінального амплітудного діапазону. Особливе призначення мають 0-й і 255-й рівні квантування. За допомогою кодів, які відповідають цим рівням, передаються сигнали синхронізації.

Аналого-цифрове перетворення сигналу яскравості описується співвідношенням

$$Y = \text{Round}(219E'_Y) + 16 ,$$

де  $E'_Y$  – аналоговий сигнал яскравості, що змінюється в діапазоні від 0 до 1 В (штрих, як це прийнято в технічній літературі щодо телебачення, означає, що сигнал пройшов гамма-корекцію);  $Y$  – цифровий сигнал яскравості, що змінюється в діапазоні від 16 до 235;  $\text{Round}(x)$  – операція округлення  $x$  до цілого числа.

Під час квантування кольорорізницевих сигналів також передбачають резервні зони – по 16 рівнів квантування зверху і знизу. На АЦП надходять не кольорорізницеві сигнали  $E'_{R-Y}$ ,  $E'_{B-Y}$ , а компресовані кольорорізницеві сигнали, що формуються відповідно до співвідношень  $E_{CR} = 0,713 \cdot E'_{R-Y}$ ,  $E_{CB} = 0,564 \cdot E'_{B-Y}$ ,

причому значення сигналів  $E_{CR}$  і  $E_{CB}$  змінюються в діапазоні від  $-0,5$  до  $0,5$  В.

Аналого-цифрові перетворення кольорорізницевих сигналів, у результаті яких утворюються цифрові кольорорізницькі сигнали  $C_R$  і  $C_B$ , виконуються відповідно до таких співвідношень:

$$C_R = \text{Round}(224E'_{CR}) + 128;$$

$$C_B = \text{Round}(224E'_{CB}) + 128.$$

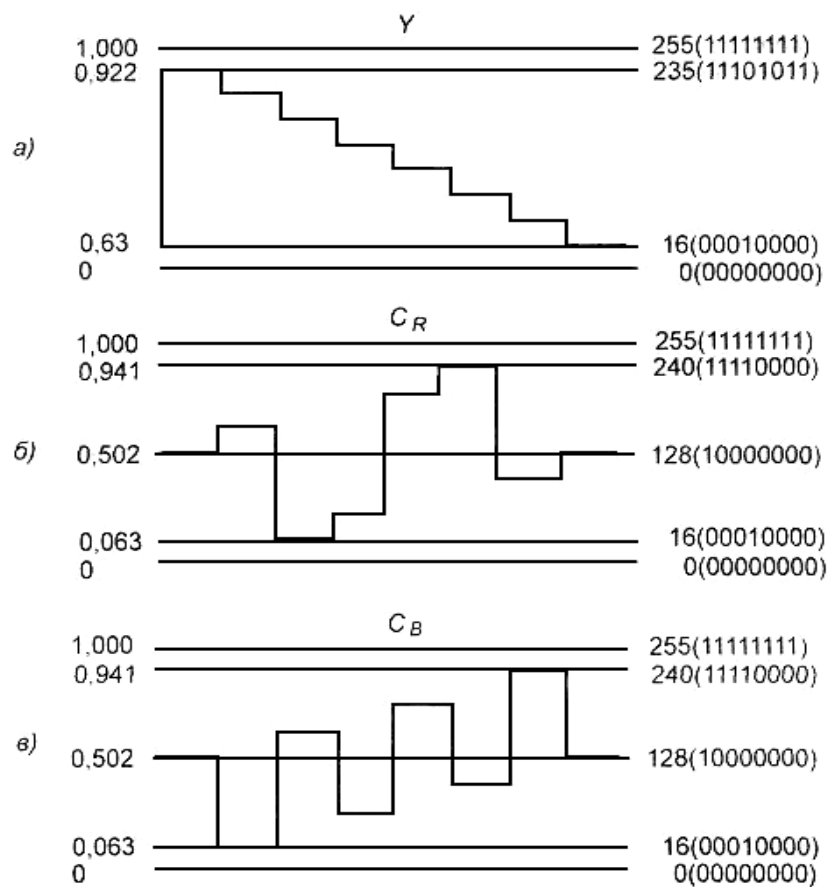
Оскільки кольорорізницькі сигнали є двополярними і максимальні відхилення від нульового значення в додатну і від'ємну сторони майже однакові, 128-й рівень квантування має відповідати нульовому значенню цих сигналів. Старший біт у цифровому кольорорізницькому сигналі показує полярність аналогового кольорорізницького сигналу.

Відповідність між рівнями аналогових телевізійних сигналів і рівнями квантування для тестового зображення у вигляді восьми кольорових смуг зображено на рис. 1.11.

До складу цифрового телевізійного сигналу згідно з Рекомендацією ITU-R BT 601 входять синхросигнали. Перед початком активної ділянки кожного рядка в кінці рядкового гасильного імпульсу передається синхросигнал початку активного рядка – ПАР (Start Active Video –SAV), а після закінчення активної ділянки кожного рядка на початку рядкового гасильного імпульсу передається синхросигнал кінця активного рядка – КАР (End Active Video – EAV).

Кожний з синхросигналів ПАР і КАР містить 4 байти. Перший байт складається з восьми двійкових одиниць, що відповідає десятковому числу 255. У разі 10-розрядного квантування замість числа 255 використовується 1023. Наступні два байти дорівнюють 0.

Призначення двійкових розрядів останнього 4-го байта пояснюється у табл. 1.1.



**Рис. 1.11.** Відповідність між рівнями аналогових телевізійних сигналів і рівнями квантування згідно з Рекомендацією ITU-R BT 601:

*а* – сигнал яскравості; *б* – кольорорізницевого червоний сигнал;  
*в* – кольорорізницевого синій сигнал



Значення  $F$  і  $V$  змінюються під час передавання сигналу КАР на початку кожного цифрового рядка. Значення перевірних бітів  $P_3...P_0$ , що залежать від значень  $F$ ,  $V$  і  $H$ , наведено у табл. 1.2. В останньому стовпці цієї таблиці подано десяткові значення числа, що передається в 4-му байті синхросигналів, яке може мати вісім допустимих значень. Перевірні біти дають можливість коректувати поодинокі і виявляти повторні помилки.

**Таблиця 1.1.**

Призначення двійкових розрядів 4-го байта  
синхросигналів ПАР і КАР

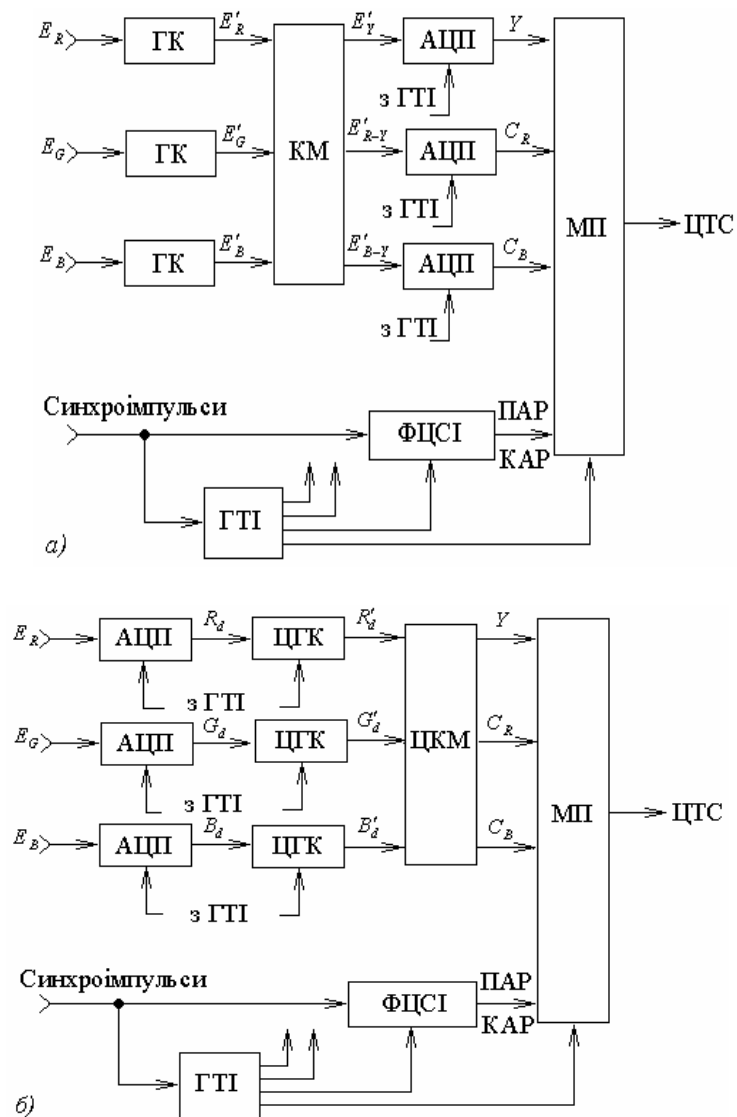
Номер розряду	Позначення	Функція
0	$P_0$	Перевірний біт
1	$P_1$	Перевірний біт
2	$P_2$	Перевірний біт
3	$P_3$	Перевірний біт
4	$H$	$H = 0$ для ПАР, $H = 1$ для КАР
5	$V$	$V = 1$ під час польового інтервалу гасіння, $V = 0$ поза цим інтервалом
6	$F$	$F = 0$ під час передавання першого поля, $F = 1$ під час передавання другого поля
7	1	Перевірний біт, постійне значення 1.

Велика частина тривалості рядкового гасильного імпульсу між синхросигналами ПАР і КАР, а саме 280 періодів тактових імпульсів з 288, залишається вільною, і в цьому інтервалі можна передавати додаткову інформацію, наприклад цифрові сигнали звукового супроводження.

**Таблиця 1.2.** Значення перевірних бітів P3...P0

$D7$	$F$	$V$	$H$	P3	P2	P1	P0	Десяткове значення
1	0	0	0	0	0	0	0	128
1	0	0	1	1	1	0	1	157
1	0	1	0	1	0	1	1	171
1	0	1	1	0	1	1	0	182
1	1	0	0	0	1	1	1	199
1	1	0	1	1	0	1	0	218
1	1	1	0	1	1	0	0	236
1	1	1	1	0	0	0	1	241

**Формувачі цифрових телевізійних сигналів.** Розглянемо два варіанти структурної схеми формувача цифрового телевізійного сигналу відповідно до Рекомендації ITU-R BT 601. У пристрої, зображеному на рис. 1.12, *a*, сигнали основних кольорів  $E_R$ ,  $E_G$ ,  $E_B$  із джерела телевізійних сигналів (телекамери) спочатку надходять на гамма-коректори (ГК). Сформовані сигнали  $E'_R$ ,  $E'_G$ ,  $E'_B$  у кодувальній матриці (КМ) за відомими співвідношеннями перетворюються в сигнал яскравості  $E'_Y$  і кольорорізницеві сигнали  $E'_{R-Y}$  і  $E'_{B-Y}$ . Далі ці сигнали перетворюються АЦП у цифрові сигнали  $Y$ ,  $C_R$  і  $C_B$  відповідно. На входах АЦП є додаткові аналогові вузли, що виконують масштабування і зсув сигналів відповідно до наведених при розгляданні квантування співвідношень. Розрядність кожного АЦП переважно дорівнює восьми.



**Рис. 1.12.** Варіанти структурної схеми формувача цифрового телевізійного сигналу:

а – з аналоговою гамма-корекцією; б – із цифровою гамма-корекцією

Синхроімпульси розгортки джерела телевізійних сигналів надходять на формувач цифрових синхроімпульсів (ФЦСІ), що виробляє синхросигнали ПАР і КАР. Крім того, синхроімпульси беруть участь у синхронізації генератора тактових імпульсів (ГТІ), який виробляє імпульси з частотами 27, 13,5 і 6,75 МГц, які надходять на інші вузли пристрою. Генератор тактових сигналів містить схему фазового автопідстроювання частоти (ФАПЧ), за допомогою якої забезпечується потрібна кількість періодів тактових імпульсів за період рядкової розгортки джерела телевізійних сигналів.

Мультиплексер (МП) у заданій послідовності передає на вихід цифрові сигнали  $Y$ ,  $C_R$  і  $C_B$  і цифрові синхросигнали, у результаті чого на виході пристрою формується цифровий телевізійний сигнал (ЦТС).

В іншому варіанті пристрою (рис. 1.12, б) сигнали основних кольорів  $E_R$ ,  $E_G$ ,  $E_B$  відразу перетворюються на цифрові сигнали  $R_d$ ,  $G_d$ ,  $B_d$  відповідно. При цьому кожен АЦП повинен мати щонайменше 10, а краще – 12 двійкових розрядів.

Далі цифрові сигнали  $R_d$ ,  $G_d$ ,  $B_d$  надходять на цифрові гамма-коректори (ЦГК), у яких відбуваються нелінійні перетворення. Кількість двійкових розрядів цифрових сигналів  $R'_d$ ,  $G'_d$ ,  $B'_d$ , які пройшли гамма-корекцію, дорівнює 8. Потім сигнали  $R'_d$ ,  $G'_d$ ,  $B'_d$  у цифровій кодувальній матриці (ЦКМ) перетворюються в цифровий сигнал яскравості  $Y$  і цифрові кольорорізницьові сигнали  $C_R$  і  $C_B$ .

Формування синхросигналів і тактових імпульсів і робота мультиплексера здійснюються аналогічно до першого варіанта пристрою. Гамма-корекція цифровими засобами забезпечує точнішу

необхідну функцію перетворення, але для цього необхідні АЦП, які мають більше двійкових розрядів, тому вони дорожчі. Нині формувачі цифрових телевізійних сигналів випускають у вигляді спеціалізованих великих інтегральних схем (ВІС), які можна використовувати, наприклад, в пристроях введення телевізійних сигналів у персональні комп'ютери.

**Паралельний відеостик.** Рекомендацією ITU-R BT 656 передбачено два варіанти інтерфейсу для передавання цифрового телевізійного сигналу: паралельний і послідовний відеостик.

Стандарт щодо паралельного відеостіку передбачає передавання цифрового телевізійного сигналу у вигляді паралельного цифрового коду. Для цього потрібно 8 (або 10) ліній зв'язку для паралельного передавання 8 (або 10) двійкових розрядів і ще одну лінію для передавання тактових імпульсів. Лінії зв'язку зазвичай виконують у вигляді скручених пар дротів, виходи передавальних пристроїв парафазні, входи приймальних пристроїв – диференційні. У такий спосіб можна передавати цифровий телевізійний сигнал усередині блоків і пристроїв або між пристроями на невеликій відстані.

У разі використання формату дискретизації 4:2:2 значення відліків яскравості  $Y$  і кольорорізницевих сигналів  $C_R$ ,  $C_B$  передаються одними і тими ж лініями в такому порядку:  $Y$ ,  $C_R$ ,  $Y$ ,  $C_B$ ,  $Y$ , ... . Частота тактових імпульсів при цьому становить  $f_T = 13,5 + 6,75 + 6,75 = 27$  МГц. Сигнали синхронізації телевізійної розгортки ПАР і КАР передаються в загальному потоці даних у відведених для них інтервалах часу.

Добуток частоти дискретизації  $f_d$  і кількості розрядів квантування  $b$  називають швидкістю передавання двійкових символів цифрового сигналу  $Q$  (біт/с). Для цифрової телевізійної

апаратури, що задовольняє вимогам таких Рекомендацій, значення цього параметра дорівнюють:

– для сигналу яскравості  $Q_Y = 13,5 \cdot 8 = 108$  Мбіт/с;

– для кольорорізницевого сигналу  $Q_C = 6,75 \cdot 8 = 54$  Мбіт/с.

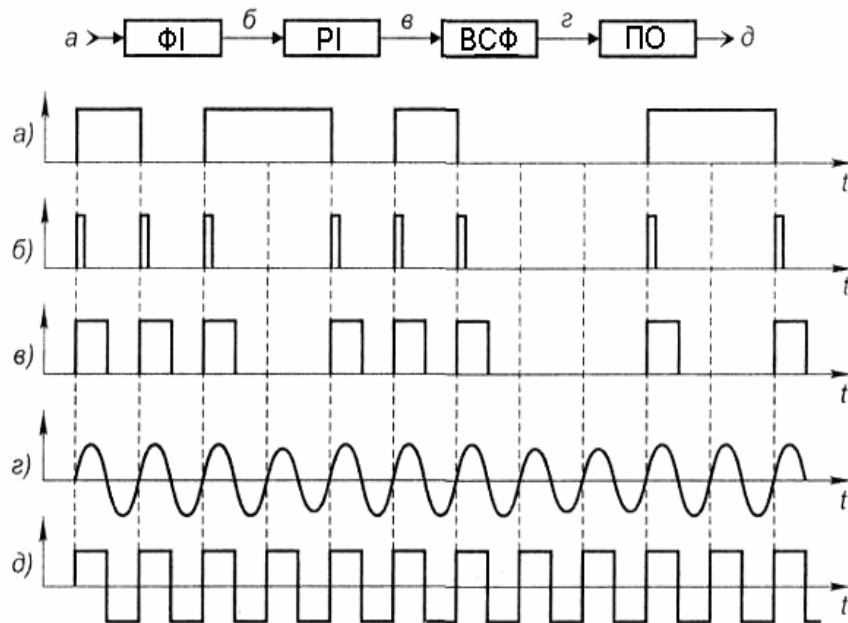
Сумарна швидкість передавання двійкових символів перетвореного у цифрову форму повного кольорового телевізійного сигналу (ПКТС) для паралельного відеостіку дорівнює  $Q_S = Q_Y + 2Q_C = 216$  Мбіт/с.

**Послідовний відеостік.** Цифровий телевізійний сигнал передається на великі відстані в послідовній формі: двійкові розряди кожного відліку цифрового сигналу передаються один за одним по одній лінії (по коаксіальному або волоконно-оптичному кабелю). Імпульси тактової частоти окремо не передаються, а відновлюються у приймальному пристрої за сигналом, який передається.

Один із простих способів відновлення тактової частоти під час приймання цифрового сигналу, що передається по послідовному каналу, проілюстровано на рис. 1.13, де точки структурної схеми і часові діаграми, що їм відповідають, позначені однаковими буквами.

Цифровий сигнал, що приймається, складається з послідовності «одиничних» і «нульових» рівнів (рис. 1.13, а). Далі сигнал надходить на формувач імпульсів (ФІ), який виробляє короткі імпульси для кожного додатного і від'ємного перепаду напруги в сигналі (рис. 1.13, б). Розширювач імпульсів (РІ) перетворює короткі імпульси в імпульси, тривалість яких дорівнює половині періоду тактової частоти (рис. 1.13, в). Ці імпульси надходять на вузькосмуговий фільтр (ВCF), який налаштований на тактову частоту. На виході фільтра виділяється синусоїдальний сигнал тактової частоти (рис. 1.13, г), який перетворюється підсилювачем-

обмежувачем (ПО) у прямокутні імпульси (рис. 1.13, д). Прямокутні імпульси використовуються для тактування сигналу, який приймається.



**Рис. 1.13.** Виділення тактових імпульсів із послідовного цифрового сигналу:

$a$  – вхідний сигнал;  $б, в, г$  та  $д$  – сигнали на виході формувача імпульсів, розширювача імпульсів, вузькосмугового фільтра та підсилювача-обмежувача відповідно

Як видно з часових діаграм (рис. 1.13), якщо в цифровому вхідному сигналі підряд ідуть декілька бітів з однаковими значеннями, то на виході формувача імпульсів протягом цього відрізка часу немає сигналів, і вихідні тактові імпульси продовжують формуватися за рахунок наявності згасального

гармонічного коливання на виході вузькосмугового фільтра. Це накладає обмеження на сигнал, який передається, оскільки передавання досить довгих послідовностей нулів або одиниць призведе до припинення формування тактових імпульсів. Крім того, на початку передавання цифрового сигналу амплітуда коливань на виході вузькосмугового фільтра наростає поступово, тому виникає деяка затримка появи тактових імпульсів на виході пристрою синхронізації.

Для подолання вказаних труднощів у системах передавання цифрових сигналів послідовними каналами зв'язку здійснюється додаткове перетворення даних (канальне кодування), у результаті якого кількість нулів або одиниць, що передаються підряд, обмежується.

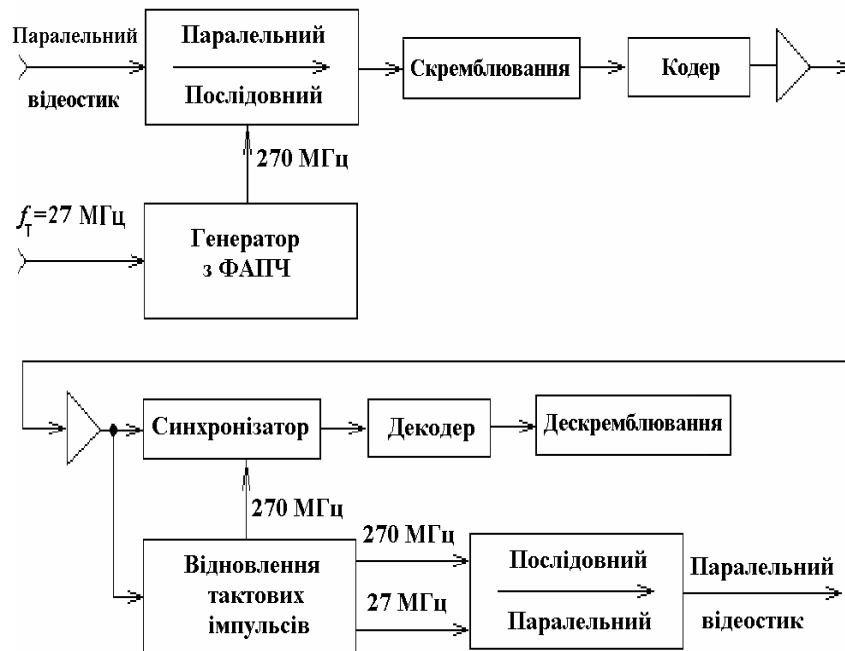
Рекомендація ITU-R BT 656 визначає параметри послідовного інтерфейсу Serial Digital Interface (SDI) для передавання цифрових телевізійних сигналів. Структурну схему системи передавання відповідно до цього інтерфейсу подано на рис. 1.14.

Цифровий сигнал, що відповідає паралельному інтерфейсу та має 10 бітів на відлік, перетворюється в послідовну форму, а тактова частота зростає у 10 разів до 270 МГц. Потім цифровий сигнал скремблюється, тобто до нього додається деякий заважаючий сигнал для захисту від несанкціонованого доступу. Після цього сигнал перетворюється в код БПН з інверсією (БПН – без повернення до нуля, англ. NRZ – Non Return to Zero).

У результаті такої обробки вилучаються довгі серії нулів і одиниць, і кожна зміна логічного рівня у сигналі на виході відповідає логічній «1» у вхідному сигналі. Це забезпечує надійне відновлення тактової частоти у приймальному пристрої і незалежність приймання від полярності вхідного сигналу.



Прийнятий сигнал декодується у звичайний послідовний двійковий код, дескремблюється, після чого в ньому відновлюються сигнали ПАР і КАР, за якими синхронізується перетворення з послідовної форми в паралельну.



**Рис. 1.14.** Послідовний відеостик

Розмах сигналу на виході передавача послідовного інтерфейсу становить близько 0,8 В, сигнал передається по коаксіальному кабелю з характеристичним опором 75 Ом. Оскільки швидкість передавання двійкових символів становить 270 Мбіт/с, ширина смуги частот кабельного каналу зв'язку має бути не меншою за 135 МГц. Реальна смуга частот коаксіальних кабелів зазвичай значно ширша.

У початковому варіанті послідовного відеостіку Рекомендацією ITU-R BT 656 передбачено передавання кожного 8-розрядного коду відліку сигналу яскравості або кольорорізницевого сигналу за допомогою 9-бітової посилки, що забезпечувало швидкість передавання двійкових символів 243 Мбіт/с.

#### **1.4. Інші стандарти цифрового подання телевізійних сигналів**

Рекомендація 601 визначає параметри цифрового подання сигналів телебачення звичайної чіткості. Для систем ТВЧ, які забезпечують значне покращення якості зображення, потрібні свої стандарти. Одним з важливих досягнень у галузі цифрового телебачення стало ухвалення Рекомендації ITU-R BT-709-3, яка визначає єдиний стандарт цифрового подання сигналів для ТВЧ, основними параметрами якого є:

- формат кадру 16:9;
- кількість активних елементів зображення в рядку 1920;
- повна кількість рядків 1125, з них активних 1080;
- частота кадрів 24, 25 або 30 Гц у разі черезрядкової або прогресивної розгортки, 50 або 60 Гц у разі прогресивної розгортки;
- частота дискретизації сигналу яскравості 74,25 МГц за частот кадрів 24, 25 або 30 Гц і 148,5 МГц за частот кадрів 50 або 60 Гц;
- частоти дискретизації кольорорізницевих сигналів у два рази менші, ніж частоти дискретизації сигналу яскравості;
- кількість розрядів квантування 8 або 10.

Новий стандарт забезпечить міжнародний обмін телепрограмами і виробництво кінофільмів за допомогою відеозапису (електронний кінематограф). Відповідно до нього ТВЧ розвивається як у Європі, так і у США.

Разом із телевізійним мовленням більш широкого поширення набуває передавання відеоінформації комп'ютерними мережами. Параметри цифрових телевізійних сигналів для систем комп'ютерного відеозв'язку визначаються Рекомендацією ІТУ-Т Н.263.

Зв'язок кількості елементів зображення у кадрі відповідно до цієї Рекомендації для формату CIF (Common Interchange Format) і похідних від нього форматів QCIF (Quarter Common Interchange Format), SQCIF (Sub-Quarter Common Interchange Format), 4CIF, 16CIF показано у табл. 1.3. В останньому стовпці таблиці подано значення максимальної швидкості передавання двійкових символів для кожного формату за частоти зміни кадрів 30 Гц.

**Таблиця 1.3.**

Параметри формату CIF і похідних від нього форматів

Формат	Сигнал яскравості	Кольорорізницеві сигнали	Максимальна швидкість передавання, Мбіт/с
SQCIF	128×96	64×48	2,9
QCIF	176×144	88×72	8,7
CIF	352×288	176×144	34,6
4CIF	704×576	352×288	139
16CIF	1408×1152	704×576	557

Слід зазначити, що у цифрових системах відеозв'язку частота зміни кадрів може бути меншою, ніж максимальне значення 30 Гц. Залежно від сюжету, який передається, ця частота переважно становить 5...15 Гц. У той же час, частота кадрів на екрані монітора

буде значно вищою (зазвичай не менше 60 Гц), оскільки в приймальній і декодувальній апаратурі здійснюються запам'ятовування прийнятих і декодованих кадрів і їх багатократне відтворення. Зменшення частоти кадрів призводить до пропорційного зменшення потрібної швидкості передавання двійкових символів у каналі зв'язку.

### **Контрольні запитання**

1. *Що таке цифрове телебачення?*
2. *Які переваги надає цифрове телебачення?*
3. *Які міжнародні організації затверджують стандарти і рекомендації у галузі цифрового телебачення?*
4. *Що таке дискретизація?*
5. *Що таке інтерполяція?*
6. *У чому проявляються Що таке дискретизація та інтерполяція? спотворення одновимірного сигналу під час дискретизації у випадку порушень умов теореми Котельникова?*
7. *Що таке просторова частота?*
8. *У яких одиницях вимірюється просторова частота у випадку неперервних зображень?*
9. *Як перетворюється просторовий спектр неперервного зображення під час дискретизації у випадках виконання й порушення умов двовимірного аналога теореми Котельникова?*
10. *Що таке квантування?*
11. *Чому похибка квантування є випадковою величиною?*
12. *Який характер розподілу імовірності похибки квантування?*
13. *Навіщо перед квантуванням виконують гамма-корекцію телевізійних сигналів?*
14. *Що означають записи «4:2:2» і «4:2:0», які характеризують*

*формат дискретизації телевізійних сигналів?*

*15. Як відбувається квантування сигналу яскравості відповідно до Рекомендації 601?*

*16. Як відбувається квантування кольорорізницевих сигналів відповідно до Рекомендації 601?*

*17. Як передаються синхроімпульси у цифровому телевізійному сигналі за Рекомендацією 601?*

*18. Що таке паралельний відеостик?*

*19. Що таке послідовний відеостик?*

*20. Розрахуйте швидкість передавання двійкових символів для цифрових телевізійних сигналів за різних форматів дискретизації.*

## 2. СТИСНЕННЯ ТЕЛЕВІЗІЙНИХ СИГНАЛІВ

### 2.1. Передавання цифрових телевізійних сигналів каналами зв'язку. Завдання стиснення відеоінформації

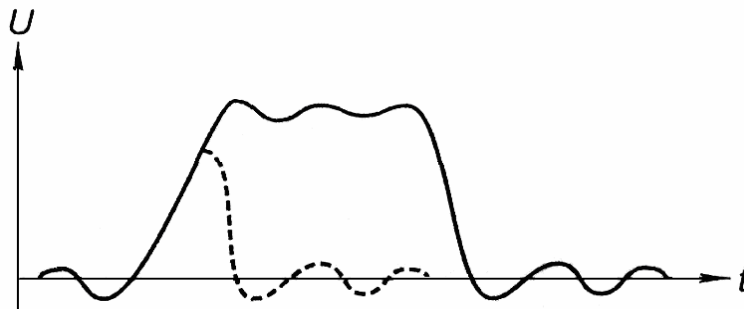
Цифровий телевізійний сигнал у послідовній формі може безпосередньо передаватися цифровими каналами зв'язку з широкою смугою частот пропускання. Для передавання цифрового телевізійного сигналу можуть використовуватися також канали зв'язку з модуляцією несучого коливання. Канали зв'язку можуть бути реалізовані як у вигляді радіоканалів, так і у вигляді коаксіальних або волоконно-оптичних кабелів. Можливість точного передавання цифрового сигналу, у першу чергу, визначається відношенням сигнал/шум (С/Ш) у каналі зв'язку.

Обчислимо потрібну для передавання цифрового телевізійного сигналу в послідовній формі ширину смуги частот каналу зв'язку. Розглянемо періодичну послідовність прямокутних імпульсів тривалістю  $T_i$ , що мають період  $T = 2T_i$ . Спектр такого сигналу містить постійну складову, основну гармоніку частоти  $1/T$  та непарні гармоніки цієї частоти, які поступово спадають.

Якщо пропустити таку послідовність через канал зв'язку з обмеженою зверху частотною смугою пропускання, то форму імпульсів буде спотворено (на рис. 2.1 показано суцільною лінією). Спотворення проявляться у вигляді збільшення тривалостей наростання і спадання імпульсу та коливальних перехідних процесів. Тривалість наростання імпульсу  $t_i$  пов'язана з верхньою граничною частотою пропускання каналу зв'язку  $f_v$  співвідношенням

$$t_i = \frac{1}{2f_v}.$$

Якщо тривалість вхідного імпульсу менша за  $t_i$ , то амплітуда імпульсу на виході каналу почне зменшуватися, як позначено на рис. 2.1 штриховою лінією. За досить високої тактової частоти та малої тривалості імпульсів, їх амплітуда на виході каналу зв'язку зрівняється з рівнем перешкод, що призведе до помилок у роботі приймального пристрою. Викиди й коливальні перехідні процеси, що виникають під час передавання імпульсів, також обмежують швидкість передавання. Коливальний перехідний процес, створений одним імпульсом, накладається на наступний імпульс. При цьому виникає перешкода, яку називають інтерференційною, що може також призвести до помилок під час приймання.



**Рис. 2.1.** Спотворення форми прямокутного імпульсу в каналі зв'язку

Для запобігання таким явищам має виконуватися умова, викладена у працях із теорії зв'язку: каналом зв'язку з шириною смуги пропускання  $\Delta F$  можна передавати  $2\Delta F$  біт інформації за 1 с, тобто ефективність використання смуги частот каналу зв'язку становить 2(біт/с)/Гц. Отже, для передавання в послідовній формі цифрового телевізійного сигналу з форматом дискретизації 4:2:2 зі швидкістю передавання двійкових символів 243 Мбіт/с ширина смуги пропускання каналу зв'язку має дорівнювати 121,5 МГц.

Зрозуміло, що ні стандартний канал наземного телевізійного мовлення з шириною смуги 8 МГц, ні супутниковий канал зв'язку з шириною смуги 27 МГц не придатні для передавання цифрового телевізійного сигналу. Це стосується і цифрового телевізійного сигналу з дискретизацією 4:2:0, і сигналів за стандартами, які застосовують у відеозв'язку (див. табл. 1.3), навіть за зниженої частоти зміни кадрів. А для передавання телевізійних сигналів високої чіткості потрібна більш широка смуга частот.

Тому однією з найважливіших завдань у галузі цифрового телебачення є зменшення швидкості передавання двійкових символів і, відповідно, потрібної смуги частот каналу зв'язку. Це завдання можна вирішити за допомогою *зменшення надмірності інформації*, яка передається в телевізійному сигналі, що зумовлює також зменшення обсягу запам'ятовувальних пристроїв під час запису телевізійних програм або окремих зображень. Відповідно до усталеної в технічній літературі термінології, зменшення обсягу переданої інформації здебільшого називають стисненням відеоінформації, а також стисненням зображень. В англійській літературі вживають термін «compression», тож замість терміна «стиснення» правомірно використовувати термін «компресія». Обернену операцію називають декомпресією (decompression) [3, 4].

У багатьох випадках замість термінів «стиснення», «компресія» або «декомпресія» будемо використовувати відповідно терміни "кодування" й "декодування", що відповідає термінології стандартів ISO (encoding, decoding). Але варто завжди розуміти, про що йдеться, оскільки у процесі стиснення за стандартами JPEG та MPEG, які розглянемо у підрозд. 2.3...2.5, виконується декілька різних по суті операцій, деякі з яких також називаються кодуванням.

Надмірність телевізійного сигналу поділяють на *структурну*,



*статистичну та психофізіологічну.*

*Структурна* надмірність пов'язана з наявністю у звичайному телевізійному сигналі гасильних імпульсів, під час яких інформація про зображення не передається. Структурну надмірність телевізійного сигналу можна зменшити за допомогою передавання під час гасильних імпульсів якої-небудь корисної інформації, наприклад, звукового супроводження, однак великого виграшу це не дає.

*Статистична* надмірність зумовлена наявністю кореляційних зв'язків між значеннями сигналу в сусідніх елементах одного рядка, у сусідніх рядках та у сусідніх кадрах. Зміст статистичної надмірності полягає у повторюваності інформації в наступних один за одним кадрах. Одним з найбільш відомих методів зменшення статистичної надмірності є кодування з передбаченням або диференціальною імпульсно-ковою модуляцією (ДІКМ). У стандартах MPEG кодування з передбаченням є одним з основних засобів стиснення. При цьому для більшості кадрів передається не саме зображення, а похибка передбачення – різниця дійсного зображення кадру й передбаченого зображення цього ж кадру, що формується за певними правилами із зображень раніше переданих кадрів. Оскільки більшість кадрів телевізійного сигналу значною мірою повторюють попередні, похибка передбачення містить значно менший обсяг інформації, ніж дійсне зображення.

*Психофізіологічна* надмірність телевізійного сигналу визначається тією інформацією, що не сприймається людиною, тож могла б і не передаватися. Таку надмірність зменшують за рахунок видалення з переданого сигналу інформації, нестача якої істотно не впливає на сприйняття зображення людиною.

Прикладом методу зменшення психофізіологічної надмірності

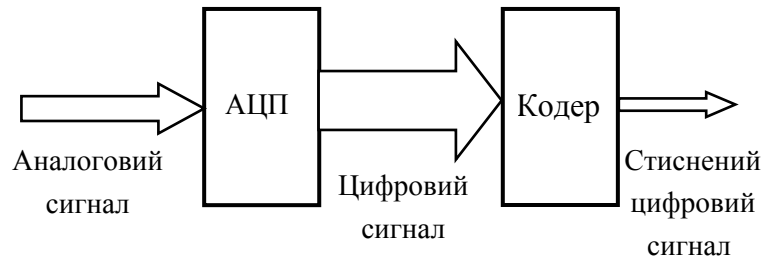
може бути кодування Кретцмера, відповідно до якого під час передавання великих деталей зображення кількість градацій яскравості збільшується, а під час передавання дрібних деталей – зменшується. Цей спосіб ґрунтується на використанні оберненої пропорційної залежності між просторовою і градаційною роздільною здатністю зорового аналізатора людини. Під час сприйняття дрібних деталей погіршується розрізнення градацій яскравості. Зазначену особливість зору використовують у методах стиснення за стандартами JPEG й MPEG, у той час як в інших методах використовують погіршення просторової та градаційної роздільної здатності зору при спостереженні за рухомими об'єктами. Спільною основою для всіх цих методів є обмеженість пропускну здатності зору як системи передавання та оброблення інформації.

Застосування різних способів стиснення інформації, яка міститься в телевізійному зображенні, дає можливість передавати стандартними телевізійними каналами з шириною смуги частот 8 МГц не тільки цифровий телевізійний сигнал звичайної чіткості, але й одночасно декількох програм телебачення звичайної чіткості, сигналів нових систем телебачення високої чіткості; а також передавати цифрові телевізійні сигнали каналами зв'язку з меншою смугою частот, ніж стандартні телевізійні канали. Співвідношення потоків інформації та потрібних значень ширини частотного діапазону пропускання каналів зв'язку проілюстровано на рис. 2.2.

Методи стиснення зображень можна поділити на дві групи: *без утрат інформації* і *з частковою втратою інформації*.

У разі стиснення без утрат після декомпресії утворюється зображення, ідентичне вхідному. Прикладами таких методів можуть бути різні алгоритми архівування зображень, які застосовуються в обчислювальній техніці й ґрунтуються на зменшенні статистичної

надмірності. Можливості стиснення дійсних кольорових або напівтонових чорно-білих зображень без утрат досить обмежені. Так, за допомогою перетворення записаного в пам'яті комп'ютера зображення у формат GIF, що є одним з найбільш ефективних, об'єм пам'яті, який займає дійсне зображення, скорочується в середньому в 2...3 рази, чого недостатньо для вирішення завдань у галузі цифрового телебачення.



**Рис. 2.2.** Співвідношення потоків інформації у різних частинах системи цифрового телебачення

Набагато більшого ефекту дають змогу досягти методи стиснення з частковою втратою інформації. Як буде зазначено у підрозд. 2.3, за допомогою методу JPEG можна зменшити обсяг інформації нерухомого зображення в 5...10 разів без помітного погіршення візуально сприйнятої якості декодованих зображень.

Можливе й ще більше стиснення, але при цьому вже спостерігаються помітні спотворення зображення. Під час стиснення телевізійних зображень за стандартами MPEG-1 і MPEG-2 обсяг інформації зменшується у 20 і більше разів завдяки зменшенню як статистичної, так і психофізіологічної надмірності.

## 2.2. Дискретне перетворення Фур'є і дискретне косинусне перетворення

Дискретне перетворення Фур'є (ДПФ) та дискретне косинусне перетворення (ДКП) є основою методів кодування стандартів JPEG, MPEG-1, MPEG-2 і деяких інших [7].

Одним з найпоширеніших способів оброблення як одновимірних, так і багатовимірних сигналів, у тому числі зображень, є подання сигналу у вигляді суми функцій з деякого набору функцій, який називають базисом.

Періодичний аналоговий сигнал  $x(t)$ , що має період  $T$ , можна подати у вигляді суми косинусоїдальних і синусоїдальних (гармонічних) коливань із частотами, кратними частоті повторення цього сигналу. Такі коливання називають гармоніками частоти, а суму – рядом Фур'є, котрий має вигляд [7]

$$x(t) = \frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos k\omega_0 t + b_k \sin k\omega_0 t), \quad (2.1)$$

де  $\omega_0 = 2\pi f_0 = 2\pi/T$  - колова частота сигналу; індекс  $k$  дорівнює номеру гармоніки частоти  $f_0$ .

У ряд Фур'є можна розкласти не тільки періодичний сигнал, але і будь-який обмежений відрізок тривалості  $T$  довільного сигналу. Початок відліку часу обирають так, щоб аналізований відрізок сигналу займав інтервал від  $-T/2$  до  $T/2$ , тоді коефіцієнти ряду Фур'є розраховують за формулами (2.2). Сума ряду Фур'є за формулою (2.1) в цьому випадку є сигналом, отриманим завдяки періодичному продовженню аналізованого відрізка з періодом  $T$  на всю вісь часу.

$$\begin{aligned}
 a_0 &= \frac{2}{T} \int_{-T/2}^{T/2} u(t) dt ; \\
 a_k &= \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos k\omega_0 t dt ; \\
 b_k &= \frac{2}{T} \int_{-T/2}^{T/2} x(t) \sin k\omega_0 t dt .
 \end{aligned} \tag{2.2}$$

Від ряду Фур'є за формулою (2.1) можна перейти до еквівалентного виразу

$$x(t) = A_0 + \sum_{k=1}^{\infty} A_k \cos(k\omega_0 t + \varphi_k), \tag{2.3}$$

де  $A_0 = a_0 / 2$ ;  $A_k = \sqrt{a_k^2 + b_k^2}$ ;  $\varphi_k = -\arctg(b_k / a_k)$ ;  $A_k$  і  $\varphi_k$  – відповідно амплітуда і початкова фаза  $k$ -ї гармоніки у ряді Фур'є. Відзначимо, що й у виразах (2.1) і (2.3) кожна частотна складова або гармоніка сигналу характеризується двома числами, оскільки має амплітуду і фазу. Амплітуди гармонік  $A_k$  утворюють спектр амплітуд сигналу, а значення початкових фаз  $\varphi_k$  утворюють спектр фаз сигналу. Розкладання в ряд Фур'є дозволяє визначити рівень різних частотних складових в сигналі.

Іншим важливим застосуванням ряду Фур'є є синтез періодичних сигналів заданої форми з гармонічних функцій. Для абсолютно точного відтворення заданого сигналу, як видно з виразу (2.1), слід підсумовувати нескінченно багато гармонік. На практиці ж беруть скінченну кількість членів ряду Фур'є, і чим більше взяти доданків, тим більш точною буде отримана форма сигналу.

У математиці і в електроніці часто використовують запис ряду Фур'є у комплексній формі

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{jk\Omega_0 t}, \quad (2.4)$$

де  $C_k$  – комплексні амплітуди гармонік;  $j = \sqrt{-1}$ .

Комплексну амплітуду можна записати у вигляді

$$C_k = C_k e^{j\varphi_k}, \quad (2.5)$$

де  $C_k$  – модуль комплексної амплітуди  $C_k$ ;  $\varphi_k$  – її фаза.

Можна показати, що

$$C_k = \frac{A_k}{2}, \quad (2.6)$$

а фаза  $\varphi_k$  та ж, що у формулі (2.3). Таким чином, комплексна амплітуда  $C_k$  містить інформацію як про амплітуду, так і про фазу відповідної гармоніки. Комплексний запис (2.4) компактний і зручний для проведення наступних математичних викладок, але він значною мірою формальний, оскільки містить доданки з від’ємними значеннями  $k$ , тобто з від’ємними частотами. Якщо підставити у формулу (2.4) вираз (2.5) для  $C_k$  і застосувати формулу Ейлера

$$e^{j\varphi} = \cos \varphi + j \sin \varphi, \quad (2.7)$$

то всі доданки з від’ємними частотами скоротяться і вийде співвідношення (2.3).

Розглянемо тепер дискретний сигнал, відмінний від 0, за  $n=0,1, \dots, N-1$ . Такий сигнал також можна подати у вигляді суми гармонічних функцій, яка міститиме скінченну кількість  $N$  доданків, кожен з яких є дискретною комплексною гармонічною функцією. Ця сума має вигляд

$$x(n) = \sqrt{\frac{1}{N}} \sum_{k=0}^{N-1} X(k) e^{j(\frac{2\pi}{N})nk}, n = 0,1,\dots, N-1, \quad (2.8)$$

де комплексні коефіцієнти  $X(k)$  визначаються співвідношеннями

$$X(k) = \sqrt{\frac{1}{N}} \sum_{n=0}^{N-1} x(n) e^{-j(\frac{2\pi}{N})nk}, k = 0, 1, \dots, N-1. \quad (2.9)$$

Обчислення коефіцієнтів за співвідношенням (2.9) називають ДПФ, а отримання сигналу за цими коефіцієнтами відповідно до виразу (2.8) – оберненим ДПФ. Частоті  $k\omega_0$ , що міститься у членах звичайного ряду Фур'є і що має розмірність рад/с, в ДПФ відповідає безрозмірна величина  $\frac{2\pi k}{N}$ , де  $k=0, 1, \dots, N-1$ .

Відношення  $\frac{k}{N}$  показує, яку частину частоти дискретизації становить частота цієї дискретної гармоніки.

Коефіцієнти ДПФ  $X(k)$  і експоненціальні множники за формулами (2.8), (2.9) є комплексними числами. Виконання ДПФ у комплексній формі призводить до значного збільшення часу обчислень. Для роботи лише з дійсними числами зазвичай використовують розкладання за допомогою ДКП. Це розкладання має вигляд

$$x(n) = \sqrt{\frac{1}{N}} \sum_{k=0}^{N-1} C(k) \cos\left(\frac{2n+1}{2N} k\pi\right), \quad (2.10)$$

де коефіцієнти ДКП обчислюють за формулами

$$C(0) = \sqrt{\frac{1}{2N}} \sum_{n=0}^{N-1} x(n),$$

$$C(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x(n) \cos\left(\frac{2n+1}{2N} k\pi\right), k = 0, 1, \dots, N-1. \quad (2.11)$$

Як і для ДПФ, обчислення коефіцієнтів  $C(k)$  за формулою (2.11)

називають прямим ДКП, а подання сигналу у вигляді суми гармонік за формулою (2.10) – оберненим ДКП.

Далі розглянемо двовимірні або просторові ДПФ і ДКП, які використовують для оброблення двовимірних сигналів, у тому числі зображень. Двовимірний дискретний сигнал, наприклад окремий кадр цифрового телевізійного сигналу, подається у вигляді матриці значень  $x(m, n)$ , де номер відліку в рядку  $m = 0, \dots, M-1$ ; номер рядка в кадрі  $n = 0, \dots, N-1$ .

Пряме двовимірне ДПФ має вигляд

$$X(k, l) = \sqrt{\frac{1}{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) e^{-j(\frac{2\pi}{M})km} e^{-j(\frac{2\pi}{N})ln}, \quad (2.12)$$

де  $k = 0, \dots, M-1$ ;  $l = 0, \dots, N-1$ .

Комплексні коефіцієнти ДПФ  $X(k, l)$  утворюють комплексний просторово-частотний спектр зображення, а модулі цих коефіцієнтів – амплітудний просторово-частотний спектр.

Обернене двовимірне ДПФ дає можливість отримати зображення у вигляді суми двовимірних дискретних гармонічних функцій

$$x(m, n) = \sqrt{\frac{1}{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} X(k, l) e^{-j(\frac{2\pi}{M})km} e^{-j(\frac{2\pi}{N})ln}. \quad (2.13)$$

Під час обчислення цієї суми залишаються тільки дійсні числа.

У наведених співвідношеннях для двовимірного ДПФ значення  $k/M$  і  $l/N$  є дискретними просторовими частотами, відповідно, за горизонтальною і вертикальною координатами. Дискретна просторова частота періодичного за горизонтальною або вертикальною координатою двовимірного дискретного сигналу дорівнює відношенню просторового періоду дискретизації до просторового періоду двовимірного сигналу по цій координаті. Просторові періоди при цьому вимірюють в одиницях відстані, а



дискретні просторові частоти є безрозмірними величинами.

Коефіцієнти двовимірного прямого ДКП визначають за формулами:

$$\begin{aligned}
 C(k, l) &= \frac{2}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) \cos\left(\frac{2m+1}{2M} \pi k\right) \cos\left(\frac{2n+1}{2N} \pi l\right); \\
 C(0, l) &= \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) \cos\left(\frac{2n+1}{2N} \pi l\right); \\
 C(k, 0) &= \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) \cos\left(\frac{2m+1}{2M} \pi k\right); \\
 C(0, 0) &= \frac{1}{2\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n).
 \end{aligned} \tag{2.14}$$

Обернене двовимірне ДКП має вигляд

$$x(m, n) = \frac{1}{\sqrt{MN}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} C(k, l) \cos\left(\frac{2m+1}{2M} \pi k\right) \cos\left(\frac{2n+1}{2N} \pi l\right). \tag{2.15}$$

Слід пам'ятати, що в різних літературних джерелах запис співвідношень для ДПФ і ДКП може бути різним. У співвідношеннях (2.8)–(2.15) нормувальні множники  $\frac{1}{\sqrt{N}}$  або  $\frac{1}{\sqrt{MN}}$  введені й у пряме, й у обернене перетворення. Іноді нормувальні множники наявні лише у співвідношеннях для прямого перетворення.

Для визначення кількості операцій, потрібних для обчислення ДПФ, перетворимо вираз (2.9) наступним чином:

$$X(k) = \sqrt{\frac{1}{N}} \sum_{n=0}^{N-1} \left[ x(n) \operatorname{Re}(W_N^{kn}) + jx(n) \operatorname{Im}(W_N^{kn}) \right], \tag{2.16}$$

де  $k = 0, 1, \dots, N-1$ ;  $W_N^{kn} = e^{j(\frac{2\pi}{N})nk}$ .

При цьому враховується, що значення сигналу  $x(n)$  є дійсними

числами.

З виразу (2.16) видно, що для обчислення одного значення  $X(k)$  потрібно виконати приблизно  $2N$  множень і  $(2N-2)$  додавань дійсних чисел; для обчислення всіх  $N$  значень  $X(k)$  слід виконати  $2N^2$  множень і  $N(2N - 2)$  додавань дійсних чисел. Крім того, потрібен запам'ятовувальний пристрій (ЗП) для зберігання значень  $x(n)$ ,  $X(k)$  і  $W_N^{kn}$ . Виконання оберненого ДПФ потребує у два рази більше операцій, оскільки значення  $X(k)$  є комплексними числами, і кількість доданків збільшиться вдвічі. У цілому можна оцінити витрати обчислювальних ресурсів під час прямого і оберненого ДПФ як пропорційні  $N^2$ .

Аналогічно можна показати, що обчислення двовимірних прямого і оберненого ДПФ вимагає виконання кількості операцій, пропорційної  $N^2M^2$ . Наприклад, обчислення ДПФ для квадратного блока зображення, що містить  $8 \times 8$  елементів (пікселів), вимагає виконання  $16 \cdot 10^3$  операцій множення і додавання. А обчислення ДПФ чорно-білого телевізійного кадру звичайного стандарту розкладання, що містить  $720 \times 576$  пікселів, вимагає виконання  $8 \cdot 10^{11}$  операцій. Якщо обчислення здійснюють на комп'ютері, що виконує 10 млн. операцій над дійсними числами за секунду, то час обчислення ДПФ становитиме близько 80 тис. секунд або більше 200 год. Очевидно, що для обчислення ДПФ телевізійних зображень у реальному часі, тобто за період кадрової розгортки, потрібно шукати способи скорочення кількості операцій.

Радикальний спосіб зменшення обсягу обчислень полягає у застосуванні відкритих у 60-ті роки минулого століття швидких алгоритмів обчислення ДПФ, названих алгоритмами швидкого перетворення Фур'є (ШПФ). Цей підхід ґрунтується на використанні

періодичності експоненціальних функцій  $W_N^{kn} = e^{j(\frac{2\pi}{N})nk}$  та їх симетрії щодо перестановки множників  $n, k$ . У найпростішому випадку, коли  $N = 2^r$ , де  $r$  – натуральне число, обчислення значення  $X(k)$  одновимірного ДПФ розбивають на  $r$  кроків. На першому кроці відліки первинного сигналу  $x(n)$  розбиваються на пари і для кожної пари обчислюються коефіцієнти ДПФ за  $N = 2$ . На другому кроці за результатами першого кроку обчислюють коефіцієнти ДПФ з  $N = 4$  для груп по 4 елементи і т. д. Аналогічно виконують обчислення оберненого ДПФ.

На кожному кроці потрібно  $N$  комплексних додавань і  $N$  комплексних множень. Кількість кроків дорівнює  $\log_2 N$ . Таким чином, кількість операцій, які потрібно виконати для обчислення прямого або оберненого ДПФ, є пропорційною  $N \log_2 N$ .

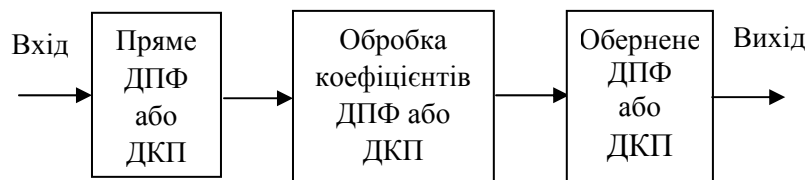
Двовимірне ШПФ може бути розкладене на послідовність одновимірних, кількість потрібних операцій пропорційна  $NM \log_2(NM)$ . Для наведеного прикладу телевізійного кадру, що містить  $720 \times 576$  пікселів, це значення дорівнює приблизно  $8 \cdot 10^6$ , що в  $10^5$  разів менше, ніж потрібна кількість операцій для безпосереднього обчислення ДПФ.

Існують також швидкі алгоритми обчислення ДКП. Як буде зазначено у підрозд. 2.3 і 2.4, у цифровому телебаченні нині основну роль відіграє ДКП блоків  $8 \times 8$  пікселів, під час якого використовується алгоритм швидкого обчислення ДКП одновимірного відрізка цифрового сигналу, що містить 8 елементів. При цьому спочатку обчислюють ДКП для кожного стовпця блока елементів зображення, а потім в одержаній матриці  $8 \times 8$  чисел обчислюються ДКП для кожного рядка. Більш детально швидкі

алгоритми обчислення ДПФ і ДКП викладено у праці [7].

У сучасних цифрових телевізійних системах ДПФ і ДКП переважно виконуються в реальному часі із застосуванням цифрових процесорів оброблення сигналів (ЦПОС), паралельних обчислювальних пристроїв та інших засобів. Такі засоби можуть входити до складу великих інтегральних схем оброблення телевізійних сигналів, наприклад, декодерів за стандартом MPEG-2.

Принцип використання ДПФ і ДКП в цифрових телевізійних системах зображено на рис. 2.3. Вхідний дискретний сигнал (одновимірний або двовимірний) піддається прямому ДПФ або ДКП, отримані коефіцієнти ДПФ або ДКП обробляються за допомогою виконання операцій над окремими частотними складовими сигналу. Наприклад, можна збільшити одні частотні складові та послабити інші. Без виконання ДПФ або ДКП таку вибірку дію на частотні складові здійснити набагато складніше. Після оброблення виконується обернене ДПФ або ДКП, внаслідок чого утворюється вихідний дискретний сигнал.



**Рис. 2.3.** Схема обробки сигналів з використанням ДПФ або ДКП

### 2.3. Стандарт JPEG

Один з найбільш ефективних і поширених методів стиснення нерухомих зображень викладено у затвердженому Міжнародною організацією стандартизації ISO стандарті JPEG (Joint Photographic

Experts Group), який визначає послідовність і параметри операцій під час кодування та декодування нерухомих зображень [8].

JPEG належить до методів стиснення зображень з утратами й використовується здебільшого для записування нерухомих зображень у запам'ятовувальних пристроях. Для більшості дійсних півтонових і кольорових зображень цей метод дає можливість зменшити обсяг інформації у 5...10 разів без помітного погіршення якості зображення, яке візуально сприймає глядач. JPEG не призначений для стиснення малюнків, креслень і зображень, що мають два рівні яскравості.

Послідовність операцій під час кодування, зображена на структурній схемі (рис. 2.4), така:

- розбивання зображення на блоки розміром  $8 \times 8$  пікселів;
- швидке двовимірне ДКП (ШДКП) у кожному блоці;
- квантування отриманих коефіцієнтів ДКП із використанням таблиці коефіцієнтів квантування (таблиці Q);
- ентропійне кодування квантованих коефіцієнтів ДКП кожного блока зображення з формуванням і використанням таблиці кодування (таблиці кодів).

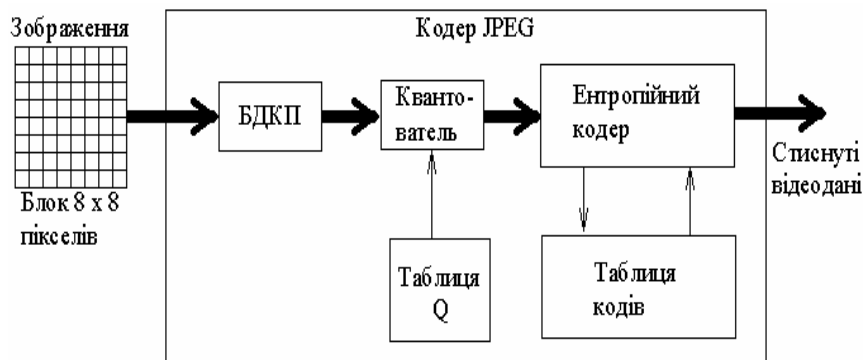
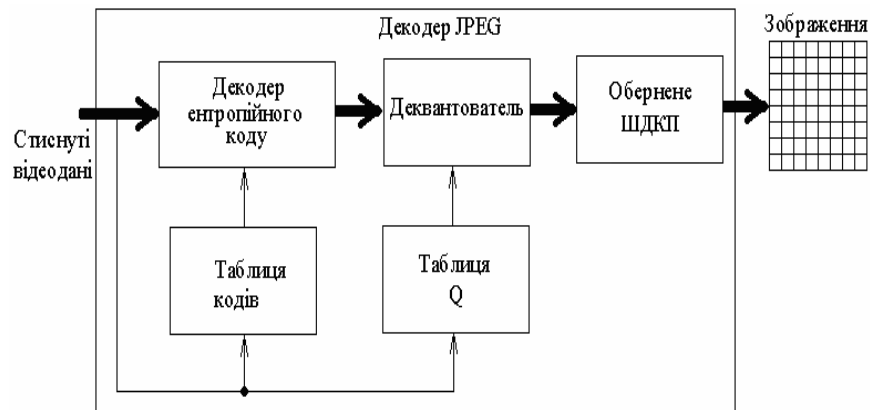


Рис. 2.4. Структурна схема кодування за стандартом JPEG

У результаті кодування вихідне зображення перетворюється у стиснені відеодані, які записуються у файл.

Послідовність операцій під час декодування, проілюстрована структурною схемою (рис. 2.5), така:

- декодування ентропійного коду;
- деквантування коефіцієнтів ДКП для кожного блока  $8 \times 8$  пікселів;
- обернене ШДКП для кожного блока;
- об'єднання блоків у декодоване зображення.



**Рис. 2.5.** Структурна схема декодування за стандартом JPEG

Для декодування ентропійного коду і деквантування використовуються відповідні таблиці, які зчитуються з файлу зі стиснутими відеоданими.

Детально розглянемо операції, які виконуються під час кодування, та відповідні обернені операції під час декодування.

**Розбивання зображення на блоки.** Півтонове монохромне (чорно-біле) зображення розбивається на блоки  $8 \times 8$  пікселів, які далі

кодуються один за одним зліва направо, один горизонтальний ряд блоків за іншим.

Якщо кількість елементів зображення в рядку не кратна 8, то праві крайні блоки в кожному горизонтальному рядку блоків доповнюються стовпцями елементів так, щоб стовпців у цих блоках було 8. За аналогією, якщо кількість рядків зображення не кратна 8, то всі блоки нижнього ряду доповнюються рядками елементів, щоб отримати 8 рядків. Стандарт рекомендує виконувати ці дії за допомогою повторення потрібну кількість разів правого крайнього стовпця або нижнього рядка. У процесі декодування всі додаткові стовпці та рядки мають вилучатися з отриманого зображення.

Кольорове зображення можна подати у форматі RGB, коли для кожного пікселя задаються значення трьох основних кольорів. У цьому випадку кожен блок  $8 \times 8$  пікселів подається трьома блоками  $8 \times 8$  чисел. Кодування даних кожного з трьох кольорів виконується так, як і для півтонового монохромного зображення.

Краще подання кольорового зображення у форматі  $YC_B C_R$ , де для кожного пікселя задаються значення сигналу яскравості та кольорорізницевих сигналів. За таких умов можливе зменшення кількості блоків для подання інформації про колір. Наприклад, якщо зменшити кількість відліків кольорорізницевих сигналів по вертикалі та по горизонталі у два рази, що відповідає формату 4:2:0, то на кожні 4 блоки елементів сигналу яскравості  $Y$  буде доводитися один блок елементів сигналу  $C_B$  та один блок елементів сигналу  $C_R$ . Порівняно з форматом RGB, повна кількість кодованих блоків зменшиться у два рази, але помітного погіршення якості зображення при цьому не відбудеться, тому що зоровий апарат людини не сприймає спотворення кольору дрібних деталей зображення. Під час об'єднання блоків у декодоване зображення

кількість елементів  $C_B$  і  $C_R$  відновлюється за допомогою інтерполяції.

**Дискретне косинусне перетворення.** Вхідні дані для ДКП мають вигляд блоків або матриць  $8 \times 8$  елементів сигналів  $Y$ ,  $C_B$  або  $C_R$ , які представлені 8- або 12-розрядними цілими додатними двійковими числами. Перед виконанням ДКП значення кожного елемента блока зміщується за рахунок віднімання числа 128 (або 2048 для 12-бітового варіанта), у результаті чого елементи блоків виражаються цілими числами зі знаком.

Після цього в кодері JPEG виконується ДКП відповідно до формул (2.11) за  $N = M = 8$ . Обернене ДКП у декодері JPEG виконується за формулою (2.10). Після нього виконується обернений зсув рівня кожного елемента за рахунок додавання числа 128 (або 2048). Зазвичай під час обчислень використовуються алгоритми швидкого ДКП.

У результаті виконання ДКП квадратної матриці  $8 \times 8$  чисел утворюється квадратна матриця  $8 \times 8$  коефіцієнтів ДКП, які можуть бути як додатними, так і від'ємними числами. Ця операція не зменшує кількості переданої інформації і є оберненою, тому що після виконання оберненого ДКП у кожному блоці та об'єднання блоків утворюється зображення, ідентичне вхідному. Єдиним джерелом необоротних утрат інформації можуть бути помилки округлення під час обчислень. Щоб уникнути цього, кількість двійкових розрядів для подання коефіцієнтів ДКП збільшується до 12, тому діапазон значень цих коефіцієнтів становить від  $-2048$  до  $2047$ .

Саме ДКП створює основу для подальшого значного зменшення обсягу інформації, що передається. Розглянемо, як це відбувається.



Передусім потрібно зауважити, що кожен коефіцієнт ДКП містить інформацію не про один якийсь елемент з матриці елементів зображення, а про всі 64 елементи. Нехай  $\{x(m, n)\}$ ,  $m, n = 0, \dots, 7$  – квадратна матриця елементів зображення, що являє собою один з його блоків,  $\{C(k, l)\}$ ,  $k, l = 0, \dots, 7$  – квадратна матриця коефіцієнтів двовимірного ДКП. Коефіцієнт  $C(0,0)$ , як випливає з формули (2.11), пропорційний постійній складовій, тобто середньому значенню  $x(m, n)$  блока зображення. Коефіцієнт  $C(0,1)$  показує значення просторово-частотної складової, що має нульову просторову частоту за горизонтальною координатою та просторовою частотою, яка дорівнює  $1/N$ , за вертикальною координатою і т. д.

У процесі оберненого ДКП відповідно до виразу (2.10) зображення формується у вигляді суми складових, кожна з яких має певну просторову частоту. Як відомо, найбільше на дійсні зображення впливають низькочастотні складові, які визначають форми і яскравість основних об'єктів і фону (див. рис. 1.6). Завдяки високочастотним складовим об'єкти набувають чітких меж і контурів, а саме зображення має дрібноструктурне заповнення (текстуру).

Можливість зменшення швидкості передавання двійкових символів за допомогою ДКП (як і за допомогою ДПФ) ґрунтується на цих властивостях просторово-частотного спектра дійсних зображень і на обмеженій здатності людського зору сприймати зміни і спотворення дрібної структури зображення. Кількість переданої інформації зменшується через більш грубе квантування частини або всіх переданих коефіцієнтів  $C(k, l)$ , у результаті якого

зменшується кількість двійкових розрядів, які використовуються для подання цих коефіцієнтів, а більшість коефіцієнтів стають рівними 0.

Як бачимо, зменшення швидкості передавання двійкових символів досягається за рахунок відкидання частини інформації. Тому зображення, відновлене за допомогою оберненого ДКП у приймальній частині системи, не буде ідентичним вхідному зображенню. Отже, такий метод кодування належить до методів кодування з частковою втратою інформації. Але інформація, яка відкидається, є несуттєвою для зорового сприйняття, тобто зміни і спотворення зображення не знижують, або майже не знижують, його суб'єктивно сприйнятої якості. Тому цей метод кодування є методом скорочення психофізіологічної надмірності зображень.

Таким чином, використання ДКП разом з подальшим квантуванням коефіцієнтів ДКП зумовлює зменшення кількості переданої інформації і, відповідно, необхідної ширини смуги частот каналу зв'язку.

**Квантування коефіцієнтів ДКП.** Квантування коефіцієнтів ДКП  $C(k,l)$  виконується відповідно до формули

$$C_q(k,l) = \text{Round} \left( \frac{C(k,l)}{Q(k,l)} \right), \quad (2.17)$$

де  $Q(k,l)$  – коефіцієнти квантування, що задають у вигляді таблиці  $8 \times 8$  цілих чисел;  $\text{Round}(x)$  – операція округлення до найближчого цілого значення;  $C_q(k,l)$  – отримані в результаті цієї операції квантування коефіцієнти ДКП, які можуть бути як додатними, так і від'ємними.

Для квантування сигналу яскравості та кольорорізницевих сигналів використовують різні таблиці. При цьому коефіцієнти

квантування  $Q(k,l)$  мають бути цілими додатними числами в діапазоні 1...255 включно. Таблиця коефіцієнтів квантування сигналу яскравості (табл. 2.1) має бути однаковою для всіх блоків певного компонента зображення і записується у файл разом зі стисненими відеоданими.

**Таблиця 2.1.** Коефіцієнти квантування сигналу яскравості

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

У результаті виконання операцій ділення та округлення багато коефіцієнтів ДКП стають рівними нулю. Саме квантування створює можливість зменшення кількості двійкових символів, потрібних для подання інформації про коефіцієнти ДКП, тобто стиснення зображення. У той же час, квантування є джерелом необоротних утрат інформації під час стиснення, для збільшення ступеня стиснення потрібно збільшувати коефіцієнти в таблицях квантування.

Операція деквантування, яка виконується в декодері JPEG, полягає у множенні коефіцієнтів  $C_q(k,l)$  на відповідні коефіцієнти  $Q(k,l)$  з таблиць квантування.

**Ентропійне кодування.** У стандарті JPEG передбачено використання двох видів ентропійного кодування: кодування за Хаффманом (Huffman) та арифметичне кодування. Кодування за Хаффманом належить до базового методу кодування і використовується підтримуватися в усіх реалізаціях JPEG. Арифметичне кодування використовується рідше й підтримується не всіма реалізаціями JPEG.

Кодування за Хаффманом відрізняється досить високою ефективністю і простотою реалізації. Щоб побудувати таблицю коду Хаффмана, що задає відповідності кодованих символів і кодових слів, потрібні лише значення частот появи символів у повідомленні, яке кодується. Ці значення частот зберігаються у файлі, тому під час декодування за ними створюється така ж таблиця, як і під час кодування. Особливістю кодів Хаффмана є те, що жодне кодове слово не є префіксом, тобто початковою частиною іншого кодового слова. Тому в закодованому повідомленні не потрібні роздільники між кодовими словами. Для оптимального кодування складових яскравості та кольору зображення потрібні різні таблиці кодів Хаффмана.

У першу чергу потрібно відзначити, що для кодування постійних складових  $C_q(0,0)$  використовують особливий метод. Якщо інші 63 коефіцієнти ДКП у кожному блоці кодуються незалежно від відповідних коефіцієнтів в інших блоках, то коефіцієнти  $C_q(0,0)$  усіх блоків кожної складової зображення попередньо кодуються з передбаченням. При цьому коефіцієнт  $C_q(0,0)$  у кожному блоці замінюється на  $DIFF = C_q(0,0) - PRED$ , де  $PRED$  – значення коефіцієнта  $C_q(0,0)$  у попередньому за порядком кодування блоці цієї ж складової. Для першого кодованого блока за  $PRED = 0$

підраховуються частоти появи різних значень  $DIFF$  у кодованому зображенні й будується таблиця коду Хаффмана, у якій значення, що зустрічаються частіше, (зазвичай ці значення малі за абсолютними величинами) подають більше короткими послідовностями двійкових символів, а значення, що зустрічаються рідше, – більш довгими. Такий метод подання значень  $DIFF$  дає змогу зменшити загальну кількість двійкових символів для подання коефіцієнтів  $C_q(0,0)$  усіх блоків зображення. Це і є ентропійне кодування.

Перед ентропійним кодуванням інших 63 квантованих коефіцієнтів ДКП у кожному блоці виконується наступна підготовча операція. Двовимірну матрицю коефіцієнтів перетворюється в одновимірну послідовність за допомогою зчитування її елементів у зигзагоподібному порядку, як зображено у табл. 2.2.

**Таблиця 2.2.** Порядок зчитування коефіцієнтів ДКП

$k/l$	0	1	2	3	4	5	6	7
0	0	1	5	6	14	15	27	28
1	2	4	7	13	16	26	29	42
2	3	8	12	17	25	30	41	43
3	9	11	18	24	31	40	44	53
4	10	19	23	32	39	45	52	54
5	20	22	33	38	46	51	55	60
6	21	34	37	47	50	56	59	61
7	35	36	48	49	57	58	62	63

По вертикалі та по горизонталі подано значення індексів  $k, l$  коефіцієнтів  $C_q(k, l)$ , у клітинках таблиці подано номери, які

отримують ці коефіцієнти в одновимірній послідовності. 0 у клітинці, що відповідає  $C_q(0,0)$ , свідчить про те, що цей коефіцієнт у цій операції не бере участі.

Після зигзагоподібного зчитування першим у послідовності буде коефіцієнт  $C_q(0, 1)$ , що відповідає низькочастотній складовій по горизонталі, потім –  $C_q(1, 0)$ , а потім усе більш і більш високочастотні складові. Послідовність завершується спеціальним символом EOB (End of Block – кінець блока).

Як відзначалося вище, у результаті квантування більшість коефіцієнтів ДКП стають рівними 0, тому в отриманій одновимірній послідовності цих коефіцієнтів утворюється велика кількість нульових елементів. Кожний відмінний від нуля коефіцієнт ДКП подається у вигляді пари чисел. Перше з цих чисел показує, скільки нульових значень підряд пройшло в послідовності перед цим ненульовим коефіцієнтом; друге число – значення самого квантованого коефіцієнта, перетворене в число зі змінною кількістю бітів. Правила цього перетворення аналогічні використуванню під час кодування постійних складових, тобто коефіцієнти з малими абсолютними значеннями подаються більш короткими послідовностями двійкових символів, а коефіцієнти з більшими абсолютними значеннями – більш довгими. Кодування за таким методом називають *runlength coding* (кодування довжини, що змінюється).

Якщо в результаті квантування утворилося багато нульових і малих за абсолютними значеннями коефіцієнтів, то таке кодування дає значний виграш: по-перше, зменшується загальна кількість чисел, що являють собою блок, який кодується; по-друге, зменшується кількість двійкових символів для подання більшості чисел.

Таким чином, для кожного блоку  $8 \times 8$  пікселів матриця квантованих коефіцієнтів ДКП перетворюється в послідовність двійкових чисел (названих відповідно до термінології теорії кодування символами), які потім піддаються ентропійному кодуванню.

Під час кодування за Хаффманом підраховуються частоти появи кожного символу в повідомленні й формується таблиця відповідності кодованих символів і кодових слів, за якою здійснюється саме кодування. Значення частот символів зберігаються у файлі. У декодері за ними формується така ж таблиця кодів, як у кодері, і за цією таблицею виконується декодування.

У процесі декодування коду Хаффмана кодові слова, які зчитуються з файлу стиснених відеоданих, перетворюються знову на послідовність чисел, за якими відновлюються значення квантованих коефіцієнтів ДКП. Важливо відзначити, що всі операції, які виконуються під час підготовки до ентропійного кодування, і саме це кодування є повністю оборотними і не спричиняють утрат інформації, а стиснення, що досягається, є наслідком раніше виконаного квантування.

Замість кодування за Хаффманом може використовуватися інший вид ентропійного кодування, названий арифметичним кодуванням. Стиснуті відеодані записуються у файл певної структури, про яку дамо лише загальні відомості.

Файл починається із заголовка, що містить різні відомості про файл; потім іде область даних про зображення, що починається з маркера SOI (Start of Image); далі можуть бути записані таблиці квантування і таблиця даних для ентропійного кодування; потім – заголовок зображення, у якому містяться відомості про розміри зображення (за кількістю пікселів), про характер зображення (чорно-

біле або кольорове), про формат дискретизації тощо. Кожний розділ файлу починається з відповідного маркера. Далі розташовано самі стиснені відеодані. Область даних завершується маркером EOI (End of Image).

Стандарт JPEG може використовуватися і для стиснення рухомих зображень, при цьому кожний кадр кодується незалежно від інших кадрів. Такий метод, який називають Motion JPEG (MJPEG), використовується для відеозапису, але він не дає достатнього ступеня стиснення відеоінформації.

## **2.4. Стандарти MPEG-1 і MPEG-2**

Стандарти стиснення рухомих зображень MPEG (Motion Picture Experts Group) розробляє і затверджує група експертів при Міжнародній організації стандартизації ISO. Стандарт MPEG-1, який здебільшого використовують під час запису відеопрограм на компакт-диски, був остаточно затверджений в 1993 році, а стандарт MPEG-2, призначений передусім для телевізійного мовлення, – у листопаді 1994 року. Стандарти MPEG-1 і MPEG-2 мають багато спільного, але між ними є й розбіжності. У цьому підрозділі переважно викладено зміст стандарту MPEG-2 і вказано його відмінності від MPEG-1 [9–11].

**Кодування й декодування рухомих зображень.** Метод кодування рухомих зображень, який використовується у стандартах MPEG-1 і MPEG-2, називають гібридним, тому що в ньому поєднано кодування всередині кадру (intraframe), спрямоване на зменшення переважно просторової надмірності в окремих кадрах, і міжкадрове (interframe) кодування, за допомогою якого зменшується надмірність, обумовлену міжкадровою кореляцією. Міжкадрове кодування дає можливість отримати істотно більший ступінь



стиснення рухомого зображення, ніж у разі роздільного стиснення окремих кадрів.

Кодування всередині кадру містить операції, багато в чому аналогічні використуваним у методі JPEG, а саме: побочне дискретне косинусне перетворення, квантування й кодування зі змінною довжиною кодових слів. Міжкадрове кодування містить такі операції, як: оцінювання та компенсація руху і кодування з передбаченням. Сутність цих операцій буде викладена у цьому підрозділі.

Цілі кадри й фрагменти кадрів можуть кодуватися із застосуванням лише кодування всередині кадру, або із застосуванням одночасно міжкадрового кодування й кодування всередині кадру (цей випадок будемо називати міжкадровим кодуванням).

**Групи зображень.** Зображенням (picture) за стандартами MPEG-1, MPEG-2 може бути як цілий кадр, так і одне з полів кадру. Далі для спрощення викладу термін «кадр» використовуватимемо замість терміна «зображення» скрізь, за винятком кадрового і польового режимів кодування.

Послідовність кадрів поділяються на групи, названі GOP (Group of picture), які містять кадри трьох типів:

- *I*-кадри (Intraframe – всередині кадру), які передаються лише із кодуванням усередині кадру та є основою для декодування інших кадрів групи, забезпечуючи можливість початку декодування і відтворення прийнятого ТВ-сигналу в будь-який момент часу;

- *P*-кадри (Predictive – передбачені), під час передавання яких виконується міжкадрове кодування за рахунок передбачення з компенсацією руху за найближчим попереднім *I*-кадром або *P*-кадром (деякі фрагменти *P*-кадру можуть кодуватися без

передбачення за допомогою кодування всередині кадру);

– *B*-кадри (Bidirectional – двонаправлені), які передаються з міжкадровим кодуванням за рахунок передбачення з компенсацією руху за найближчими до них як попереду, так і позаду *I*-кадрами і *P*-кадрами, а самі не можуть використовуватися для передбачення інших кадрів (деякі фрагменти *B*-кадру можуть передаватися з кодуванням усередині кадру).

Розглянемо приклад послідовності кадрів:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
I	B	B	P	B	B	P	B	B	P	B	B	P	B	B	I	B	B	P	B

Кадри з 1 по 15 утворюють групу кадрів, кількість кадрів у якій може бути й іншою, але група завжди починається з *I*-кадру. *P*-кадр 4 передбачається за *I*-кадром 1, *P*-кадр 7 – за *P*-кадром 4, *P*-кадр 10 – за *P*-кадром 7 і т. д. *I*-кадр 16 передається з кодуванням усередині кадру, незалежно від усіх попередніх кадрів. *B*-кадри 2 і 3 передбачаються за *I*-кадром 1 і за *P*-кадром 4, *B*-кадри 5 і 6 – за *P*-кадрами 4 і 7 і т. д. *B*-кадри 14 і 15 передбачаються за *I*-кадром 16 і за *P*-кадром 13. Оскільки кожний *B*-кадр повинен іти після обох кадрів, за якими він передбачається, порядок проходження кадрів змінюється таким чином:

1	4	2	3	7	5	6	10	8	9	13	11	12	16	14	15	19	17	18	22
I	P	B	B	P	B	B	P	B	B	P	B	B	I	B	B	P	B	B	P

Для переставлення кадрів потрібен запам'ятовувальний пристрій обсягом мінімум на 4 кадри без стиснення. У такому порядку кадри кодуються й передаються, а у процесі декодування відновлюється початковий порядок кадрів, для чого теж потрібний запам'ятовувальний пристрій на декілька кадрів.

**Макроблоки.** *Макроблоком* називають квадратний фрагмент зображення розміром  $16 \times 16$  елементів (пікселів), що містить інформацію як про яскравість ( $Y$ ), так і про колір ( $C_B$ ,  $C_R$ ). У разі використання формату дискретизації 4:2:0 кожний макроблок містить 4 блоки  $8 \times 8$  елементів сигналу яскравості  $Y$  і по одному блоку  $8 \times 8$  елементів кольорорізницевих сигналів  $C_B$  і  $C_R$ ; за формату 4:2:2 кожний макроблок містить по два блоки  $C_R$  і  $C_B$ , а за формату 4:4:4 – по чотири блоки  $C_R$  і  $C_B$ .

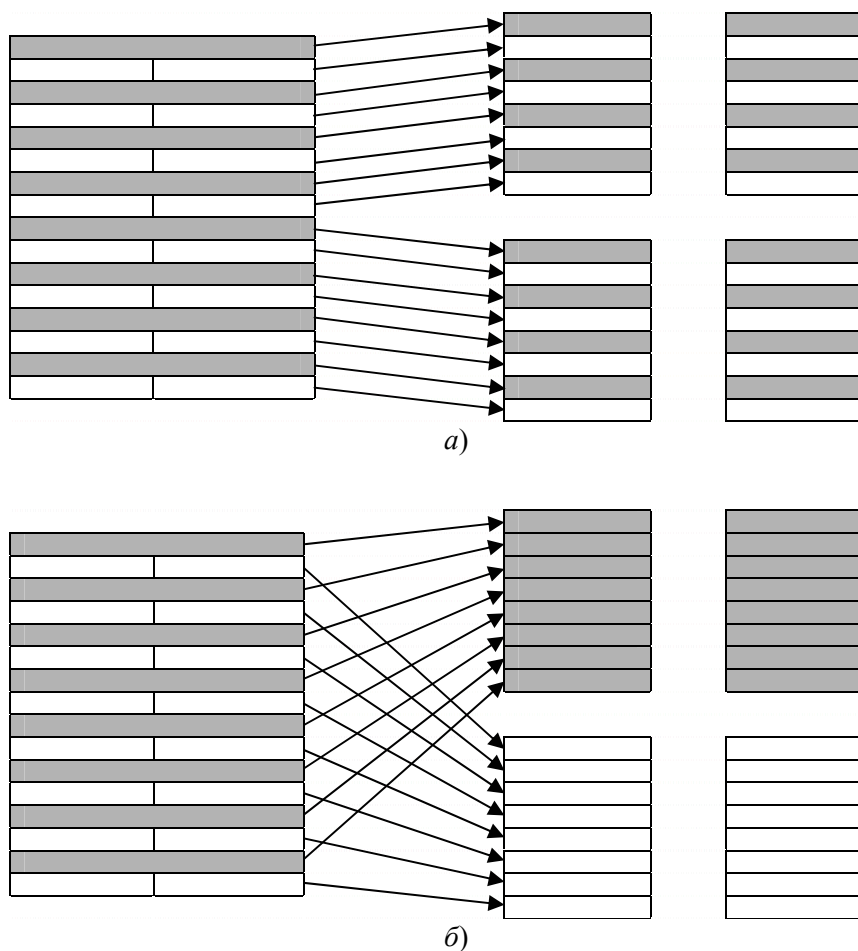
Групу наступних один за одним макроблоків називають *слайсом* (slice – частка, частина, квант). Кількість макроблоків у слайсі може бути довільною. Слайси у зображенні не повинні перекриватися, але вони можуть переміщуватися від одного зображення до іншого. У межах слайса зберігаються постійні параметри стиснення. Крім того, заголовки слайсів забезпечують відновлення правильного декодування після збою.

**Прогресивна й черезрядкова розгортки.** Під час кодування телевізійного зображення, яке передається з прогресивною розгорткою, кожний кадр складається з одного поля і розбивається на макроблоки.

У разі черезрядкової розгортки кожний кадр складається з двох полів: перше містить непарні рядки кадру, а друге – парні рядки. При цьому можливі два варіанти кодування кадру, вибір одного з яких здійснюється за результатами оцінки руху в ньому.

У кадровому режимі кодування (frame) кодованим зображенням є повний кадр, що цілком зберігається в запам'ятовувальному пристрої кодера, внаслідок чого для кодування доступні одночасно як парні, так і непарні рядки. Блоки елементів сигналу яскравості та кольорорізницевих сигналів для форматів 4:2:2 і 4:4:4 виділяються з макроблока, як зображено на рис. 2.6, а, де темними показано

непарні рядки, а світлими – парні. За формату 4:2:0 у блоки кольорорізницевих сигналів входять елементи з кожного другого рядка. Кадрове кодування застосовується у випадках, коли зміни у другому полі кадру щодо першого поля того ж кадру незначні.



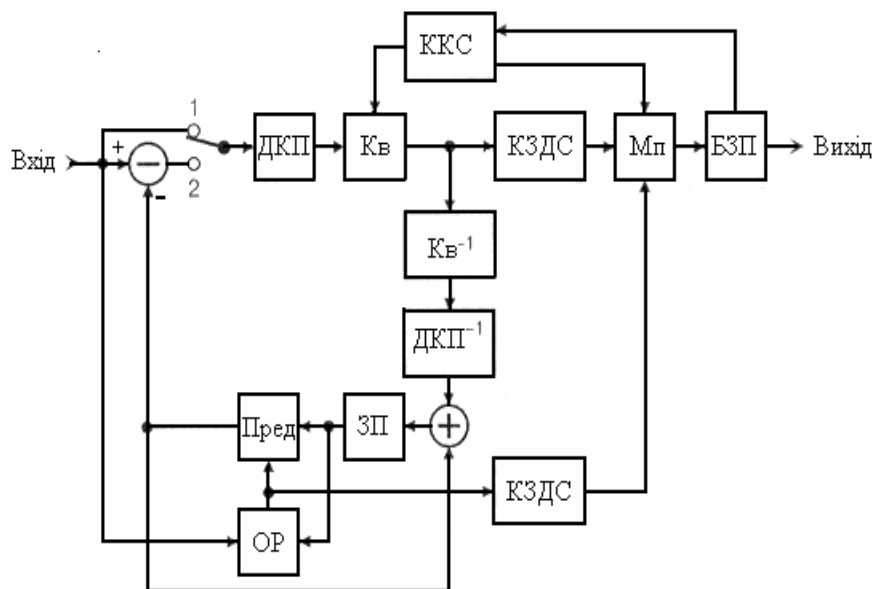
**Рис. 2.6.** Формування блоків у разі черезрядкової розгортки за кадрового (а) і польового (б) режимів

У польовому режимі кодування (field) кодується кожне поле, причому перше поле кадру може використовуватися для передбачення макроблоків другого поля того ж кадру. При цьому в кожний блок елементів сигналу яскравості або кольорорізницевих сигналів для форматів 4:2:2 і 4:4:4 входять елементи з одного поля, як показано на рис. 2.6, б. Блоки елементів сигналів для формату 4:2:0 утворюються так само, як за кадрового кодування.

**Кодер відеоінформації.** У стандартах MPEG не описано структуру кодера, а лише визначено параметри потоку даних на його виході. Структурна схема кодера відеоінформації (рис. 2.7) відображає основні операції, які виконуються під час кодування і забезпечують отримання вихідного потоку даних з потрібними параметрами.

На схемі позначено: ДКП – блок прямого дискретного косинусного перетворення;  $\text{ДКП}^{-1}$  – блок оберненого дискретного косинусного перетворення; Кв – квантувач;  $\text{Кв}^{-1}$  – деквантувач, тобто блок, що виконує обернену квантуванню операцію; ЗП – запам'ятовувальний пристрій; Пред – блок, що формує передбачений кадр; ОР – блок оцінки руху і формування векторів руху; КЗДС – блок, у якому виконується кодування зі змінною довжиною кодового слова; Мп – мультиплексер; БЗП – буферний запам'ятовувальний пристрій; ККС – блок керування коефіцієнтом стиснення зображення. Крім того, на схемі показано пристрої, які виконують додавання та віднімання, а також перемикач. Робота всіх блоків синхронізується загальною тактовою частотою 27 МГц.

У кодері реалізуються два режими кодування: кодування всередині кадру (перемикач у положенні 1) і міжкадрове кодування з передбаченням і компенсацією руху (перемикач у положенні 2).



**Рис. 2.7.** Структурна схема відеокодера MPEG-2

Всі макроблоки *I*-кадрів кодуються в режимі кодування всередині кадру. Операції при цьому аналогічні операціям стандарту JPEG: розкладання на блоки  $8 \times 8$  пікселів, поблочне ДКП, квантування отриманих коефіцієнтів, зчитування в зигзагоподібному порядку, кодування зі змінною довжиною кодових слів.

Квантування виконується відповідно до співвідношення

$$C_q(k, l) = \text{Round} \left( \frac{C(k, l)}{fQ(k, l)} \right), \quad (2.18)$$

де використано ті ж позначення, що й у виразі (2.17), а  $f$  - масштабний коефіцієнт. Під час квантування використовуються таблиці, аналогічні табл. 2.1. У разі використання нестандартних таблиць коефіцієнтів квантування, вони включаються в загальний потік даних на виході. Ступінь стиснення регулюють масштабним

коефіцієнтом, що набуває значень від 1 до 31 і може змінюватися у процесі передавання. Для постійних складових  $C_q(0, 0)$  коефіцієнт квантування дорівнює 8 і не залежить від масштабного коефіцієнта.

Кодування зі змінною довжиною кодових слів здійснюється за допомогою таблиць кодів, заданих у стандарті. Як і в JPEG, коефіцієнти ДКП, що відповідають постійним складовим, кодуються з використанням передбачення за такими ж коефіцієнтами попередніх блоків, і для них призначені окремі таблиці кодів зі змінною довжиною кодових слів. Окремі таблиці кодування передбачено для різних видів службової інформації.

Макроблоки  $P$ -кадрів можуть кодуватися як методом усередині кадру, так і міжкадровим, що ґрунтується на передбаченні і компенсації руху, і під час якого передаються кодовані похибки передбачення. Зображення, за яким виконується передбачення, формується з кодованих, а потім декодованих даних попереднього  $I$ -або  $P$ -кадру. У деквантувачі дані помножуються на коефіцієнти квантування, потім виконується обернене ДКП так само, як у декодері приймальної частини системи, після чого декодоване зображення записується у ЗП. Таким чином, для передбачення використовуються зображення, що містять похибки, внесені квантувачем. Ці похибки потрапляють у передбачені зображення, а під час формування різниці дійсного та передбаченого зображень – знову в кодер, але вже з протилежним знаком. У результаті чого утворюється від'ємний зворотний зв'язок, що допомагає уникнути накопичення похибок квантування. Для реалізації всіх можливих варіантів передбачення ЗП має містити декілька (як мінімум 4) попередніх кадрів.

Під час формування передбаченого макроблока використовується оцінка й компенсація руху методом *відповідності блоків* (Block

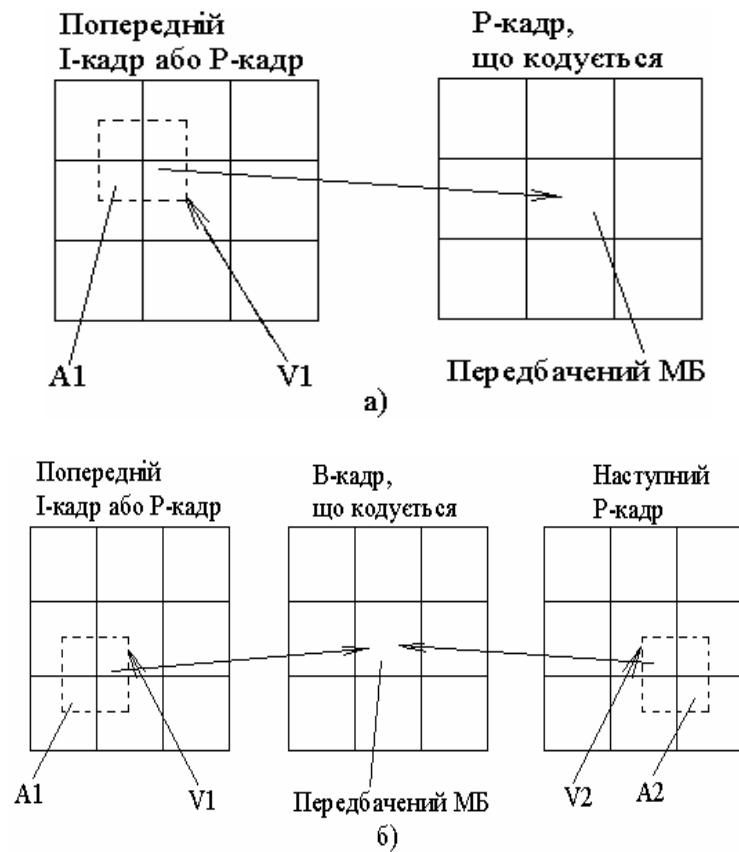
Matching). Оцінка руху здійснюється через порівняння кодованого зображення (кадру або поля), що надходить на вхід кодера, з раніше переданим зображенням, що перебуває у ЗП, і використовується для передбачення (*опорним зображенням*) (рис. 2.8,а). Для кожного макроблока кодованого зображення відшукується відповідна йому, тобто найбільш схожа на нього ділянка A1 розміром  $16 \times 16$  елементів в опорному зображенні. При цьому послідовно перебираються вектори руху (або вектори зсуву)  $(\Delta n, \Delta m)$ , де  $\Delta n, \Delta m$  – збільшення координат по горизонталі й по вертикалі, що набувають значень у деяких заданих інтервалах. Для кожного вектора руху використовують ділянку  $16 \times 16$  пікселів в опорному зображенні, зміщену на цей вектор відносно аналізованого макроблока кодованого зображення, і розраховують суму  $S(\Delta n, \Delta m)$  абсолютних значень різниць елементів макроблока кодованого зображення і зміщеної ділянки опорного зображення

$$S(\Delta n, \Delta m) = \sum_{n, m \in B} (x'(n + \Delta n, m + \Delta m) - x(n, m)), \quad (2.19)$$

де  $x(m, n)$  – елемент першого кадру;  $x'(m, n)$  - елемент другого кадру;  $m, n$  - дискретні координати по вертикалі й по горизонталі, відлік яких виконується, наприклад, від лівого верхнього кута макроблока кодованого зображення. Підсумовування здійснюється за всіма елементами макроблока. Оцінка руху здійснюється тільки за елементами сигналу яскравості.

З усіх перевірених векторів  $(\Delta n, \Delta m)$  обирається той, який забезпечує найменше значення суми за виразом (2.19) і вважається далі вектором руху для такого макроблока. У деяких випадках замість суми абсолютних значень використовується корінь із суми квадратів різниць, однак його розрахунок вимагає більшого обсягу обчислень.





**Рис. 2.8.** Отримання передбаченого макроблоку в *P*-кадрі (а) і у *B*-кадрі (б)

Повний перебір можливих зсувів у заданому діапазоні забезпечує обчислення абсолютного мінімуму суми за формулою (2.19), але вимагає виконання  $(2W + 1)^2$  обчислень за цією формулою, де  $W$  – максимальний зсув за якою-небудь координатою, виражений у кількості пікселів. Є різні алгоритми (або стратегії) пошуку відповідності блоків зі зменшеним обсягом обчислень.

Наприклад, спочатку виконується пошук вектора руху з більшим кроком зсуву (у декілька пікселів) по всьому діапазону зсувів, а потім у зоні, де виявлено мінімум суми за виразом (2.19), виконується пошук із кроком 1 або навіть  $1/2$  піксела для точного визначення вектора руху. У результаті положення відповідної ділянки визначається з точністю до половини піксела за обома координатами.

Якщо під час пошуку не знайдено відповідної ділянки, відмінність якої від цього макроблока (значення  $S$ ) не перевищує встановленого порогового значення, то цей макроблок кодується у режимі всередині кадру за аналогією до макроблоків  $I$ -кадрів.

Якщо відповідну ділянку  $A1$  знайдено, то макроблок кодується в міжкадровому режимі, і для нього визначається вектор руху  $V1$ . Вектори руху кодуються зі змінною довжиною кодових слів за допомогою особливих таблиць кодування, заданих у стандарті, і через мультиплексер включаються в загальний потік даних.

Якщо для макроблока обрано міжкадровий режим кодування, то формується передбачений макроблок (МБ), у якості якого береться знайдена відповідна ділянка  $A1$  з опорного зображення. Це і є компенсація руху. Передбачений макроблок поелементно віднімається зі справжнього макроблока. Отриманий різницевий макроблок (похибка передбачення) піддається поблочному ДКП, квантуванню й кодуванню зі змінною довжиною кодових слів. Матриця коефіцієнтів квантування для похибок передбачення містить 64 числа і використовується для квантування складової яскравості і кольорорізницевих складових.

Для макроблоків  $B$ -кадрів пошук відповідної ділянки здійснюється як у попередньому  $I$ - або  $P$ -кадрі, так і у наступному  $P$ -кадрі (рис. 2.8, б), і залежно від результатів пошуку можливі такі

варіанти кодування макроблоків:

- макроблок кодується у режимі всередині кадру;
- формується передбачений макроблок у вигляді відповідної ділянки  $A1$  попереднього  $I$ - або  $P$ -кадру;
- формується передбачений макроблок у вигляді відповідної ділянки  $A2$  наступного  $P$ -кадру;
- формується передбачений макроблок у вигляді поелементної напівсуми відповідної ділянки  $A1$  попереднього  $I$ - або  $P$ -кадру і відповідної ділянки  $A2$  наступного  $P$ -кадру, тобто у вигляді результату інтерполяції за цими ділянками.

В останніх трьох варіантах макроблок кодується в міжкадровому режимі, як це було описано для  $P$ -кадрів. У випадку передбачення за рахунок інтерполяції для макроблока потрібно передавати 2 вектори руху  $V1$  і  $V2$ , що показують положення відповідних ділянок у попередньому та у наступному кадрах.

Під час кодування телевізійних зображень з черезрядковою розгорткою можливі два варіанти передбачення: польовий і кадровий. У разі польового режиму макроблоки кожного поля передбачаються незалежно від іншого поля цього кадру, і для формування макроблока використовуються дані одного або двох раніше кодованих полів. У разі кадрового режиму для формування передбаченого макроблока використовують дані, що містяться в обох полях одного або двох раніше кодованих кадрів. Доступні також додаткові режими передбачення у разі черезрядкової розгортки.

Крім того, стандарт MPEG-2 дозволяє пропускати деякі макроблоки під час кодування (Skipped macroblock), для яких ніякі дані не передаються. Такий варіант використовується, якщо макроблок, що кодується, не відрізняється від відповідного

макроблока в опорному зображенні.

Описаний спосіб кодування телевізійного сигналу називають кодуванням із передбаченням та компенсацією руху. Виграш у стисненні зображення досягається завдяки тому, що різниці дійсних і передбачених макроблоків *P*- і *B*-кадрів містять значно менше інформації, ніж самі ці макроблоки. При цьому для *B*-кадрів обсяг переданої інформації буде найменшим, тому що у разі двоспрямованого передбачення похибка передбачення мінімальна.

Кодовані відеодані та вектори руху через мультіплексер надходять у БЗП (див. рис. 2.7), що працює за принципом «першим увійшов – першим вийшов». Одна з функцій БЗП – узгодження нерівномірного в часі потоку даних після кодування із суворо постійною швидкістю передавання двійкових символів у процесі кодування в реальному часі. Нерівномірність потоку даних, що надходять у БЗП, зумовлена, у першу чергу, наявністю різних типів кадрів. Зчитування даних із БЗП здійснюється з постійною швидкістю. Ступінь заповнювання БЗП коливається в часі, зростаючи у разі збільшення потоку даних, які надходять на нього, і знижуючись у разі зменшення цього потоку.

На швидкість заповнення буфера впливає також характер зображення, що передається. Якщо в зображенні багато дрібних деталей, то зростають кількість і рівні високочастотних складових просторово-частотного спектра, тобто кількість відмінних від нуля коефіцієнтів ДКП. Це призводить до збільшення потоку даних. У разі ж передавання «гладких» зображень кількість відмінних від нуля коефіцієнтів ДКП зменшується, тому що просторово-частотний спектр зображення має здебільшого низькочастотні складові.

Для оптимізації роботи системи бажано підтримувати рівень заповнення БЗП майже постійним, оскільки якщо БЗП

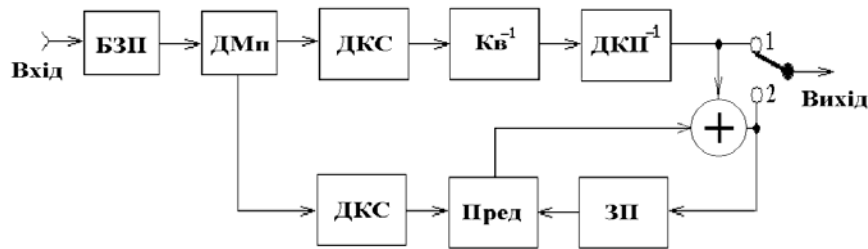
переповнюється, то буде втрачатися частина даних, що зумовить погіршення якості зображення на виході системи. Якщо ж БЗП повністю звільняється, то каналом зв'язку доводиться передавати «порожні» блоки (stuffing), що призводить до зниження ефективності його використання. Щоб уникнути обох небажаних наслідків, у кодері зображення введено зворотний зв'язок із БЗП на блок ККС, який керує ступенем стиснення зображення.

Суть дії цього зворотного зв'язку полягає в тому, що якщо передається зображення з дрібною структурою і заповнення БЗП збільшується, то під впливом зворотного зв'язку збільшується масштабний коефіцієнт квантування (див. формулу (2.18)). При цьому кількість бітів на кожний коефіцієнт ДКП зменшується, і значення потоку даних підтримується майже постійним. І навпаки, у разі передавання «гладких» зображень квантування стає більш точним. Такий метод відповідає властивостям людського зору: на зображеннях з дрібною структурою менше помітні спотворення під час передавання рівнів яскравості, тому що в першу чергу сприймаються контури деталей. Параметр квантування може змінюватися або після кодування кожного кадру з урахуванням його типу, або в межах одного кадру після кодування кожного слайсу. Дані про параметр квантування включаються в загальний потік даних у заголовках слайсів.

**Декодер відеоінформації.** Відповідно до стандарту у відеодекодері (рис. 2.9) виконуються декодування кодів змінної довжини, деквантування, обернене ДКП, компенсація руху і відновлюється вхідна послідовність кадрів.

Декодер містить буферний запам'ятовувальний пристрій (БЗП); демультіплексор (ДМп); декодери кодів зі змінною довжиною кодових слів ДКС, а також деквантувач  $K_v^{-1}$ , блок зворотного

дискретного косинусного перетворення ДКП<sup>-1</sup> і інші блоки, аналогічні відповідним блокам кодера. Тактова частота 27 МГц відновлюється з використанням даних з декодованого потоку.



**Рис. 2.9.** Структурна схема відеодекодера MPEG-2

На вході декодера БЗП виконує функцію узгодження постійної швидкості передавання двійкових символів у вхідному потоці даних із процесами в декодері, за яких дані з БЗП зчитуються нерівномірно у часі. З виходів демультимплексора кодовані дані зображення та значення параметра квантування надходять на ДКС і далі на деквантувач, а вектори руху надходять на ДКС і далі на блок Пред, що формує передбачений кадр.

Так само як і в кодері, у декодері є два режими роботи. Під час приймання *I*-кадрів і макроблоків *P*-кадрів і *B*-кадрів, які передаються з кодуванням всередині кадру, на виході блоку оберненого ДКП (ДКП<sup>-1</sup>) формуються блоки зображення. Перемикач на структурній схемі при цьому перебуває в положенні 1, і сигнал з блоку ДКП<sup>-1</sup> спрямовується на вихід. Під час приймання макроблоків *P*-кадрів і *B*-кадрів, які передаються з кодуванням у міжкадровому режимі, перемикач перебуває в положенні 2. У цьому випадку формування вихідного сигналу відбувається за рахунок поелементного додавання значень різниць, які надходять із блоку ДКП<sup>-1</sup>, із передбаченим макроблоком, сформованим з елементів

раніше декодованих зображень із використанням декодованих векторів руху.

Реалізація декодера апаратними, програмними або апаратно-програмними засобами істотно простіша, ніж реалізація кодера, тому що в декодері не треба виконувати пошук відповідних ділянок у зображеннях, за якими виконується передбачення, а саме цей пошук вимагає найбільшого обсягу обчислень.

**Масштабованість.** Важливою особливістю стандарту MPEG-2 є *масштабованість* (Scalability), яку визначають як можливість отримання зображення з частини повного потоку відеоданих, який за наявності масштабованості складається з двох або більше шарів (layers). Базовий шар дає змогу отримати зображення з деякими початковими параметрами якості, а додаткові або поліпшуючі (enhancement) шари потоку даних – отримати зображення поліпшеної якості. Стандарт MPEG-2 передбачає організацію потоків відеоданих як із масштабованістю, так і без неї.

Стандартом MPEG-2 передбачено такі види масштабованості, як: за просторовою роздільною здатністю (Spatial Scalable), за відношенням сигнал/шум (SNR Scalable), за часом (Temporal Scalable) і за розділенням даних (Data partitioning Scalable). Кожний вид масштабованості, взятий окремо, припускає наявність у потоці даних двох рівнів. У разі використання одночасно двох видів масштабованості кількість рівнів у потоці даних не більше трьох.

*Масштабованість за просторовою роздільною здатністю* полягає в отриманні від одного джерела відеоінформації двох телевізійних сигналів з різними параметрами за роздільною здатністю, наприклад телевізійних сигналів звичайної чіткості та високої чіткості. Базовий шар потоку даних містить інформацію, достатню для відтворення зображення звичайної чіткості.

Додатковий шар містить дані, що дають змогу доповнити відтворене зображення до зображення високої чіткості. Важливо відзначити, що обсяг цих додаткових даних менший, ніж повний обсяг даних про зображення високої чіткості, тому що частина інформації передається в базовому шарі.

Користувачі, які мають більш складні і дорогі декодери з можливістю декодування обох шарів потоку даних, отримуватимуть на екранах своїх приймачів зображення високої чіткості. Інші користувачі, які мають більш прості та дешеві декодери для декодування лише базового шару, також зможуть дивитися ці передачі, але у вигляді зображення звичайної чіткості.

*Масштабованість за відношенням сигнал/шум (С/Ш)* дає можливість отримувати від одного джерела відеоінформації зображення з двома рівнями відношення сигнал/шум, а, отже, з двома рівнями якості. Під шумом розуміють похибки, внесені квантуванням і кодуванням. Базовий шар потоку даних може містити зображення з більшим ступенем стиснення, тобто менш якісне. Додатковий шар при цьому міститиме дані, що допомагають у разі додавання їх до даних базового шару отримати зображення з меншим ступенем стиснення, тобто більш якісне.

*Масштабованість за часом* дає можливість отримувати від одного джерела відеоінформації телевізійні зображення з двома рівнями роздільної здатності за часом, наприклад, базовий шар може містити звичайний телевізійний сигнал із частотою кадрів 25 Гц і черезрядковою розгорткою, а додатковий шар – дані, які допомагають у разі додавання їх до даних базового шару отримати телевізійне зображення з частотою кадрів 50 Гц і прогресивною розгорткою.

*Масштабованість за розділенням даних* дає змогу



використовувати для передавання паралельно два канали зв'язку. Одним каналом, більш захищеним від перешкод, передається базовий шар, що містить найбільш критичну до похибок інформацію: заголовки, вектори руху, коефіцієнти ДКП, які відповідають низьким просторовим частотам. Менш захищеним каналом передаються менш критичні до похибок дані, наприклад, коефіцієнти ДКП, що відповідають високим просторовим частотам.

Цей вид масштабованості добре узгоджується з попередніми трьома видами, оскільки під час використання кожного з них базовий шар потоку даних може передаватися більш захищеним від перешкод каналом зв'язку, а додатковий шар – менш захищеним. Тоді за хороших умов приймання користувач, що має декодер для обох шарів, зможе бачити зображення найвищої якості, а у разі погіршення цих умов він отримуватиме менш якісне зображення.

Слід зазначити, що масштабованість є важливою передумовою подальшого розвитку цифрових телевізійних систем. Підхід, який ґрунтується на масштабованості, в останні роки став характерним не лише для цифрового телебачення, але і для багатьох інших телекомунікаційних і інформаційних технологій.

**Рівні та профілі MPEG-2.** Різні варіанти телевізійних систем і методів кодування телевізійних сигналів, передбачених стандартом MPEG-2, подано у табл. 2.3. Чотири рядки таблиці відповідають чотирьом рівням просторової роздільної здатності:

- Low (352×280 елементів) – рівень телебачення зниженої чіткості, який використовують у відеотелефонах і техніці телеконференцій;
- Main (720×576 елементів) – рівень телебачення звичайної чіткості;
- High-1440 (1440×1152 елементів) – рівень телебачення високої

чіткості з форматом зображення 4:3;

– High (1920×1152 елементів) – рівень телебачення високої чіткості з форматом зображення 16:9.

**Таблиця 2.3.** Рівні та профілі стандарту MPEG-2

Рівень	Профіль				
	Простий (Simple)	Основний (Main)	Масштабування за С/Ш (SNR Scalable)	Просторовий (Spatially Scalable)	Високий (High)
High 1920×1152	–	80 Мбіт/с	–	–	100 Мбіт/с
High-1440 1440×1152	–	60 Мбіт/с	–	60 Мбіт/с	80 Мбіт/с
Main 720×576	15 Мбіт/с	15 Мбіт/с	15 Мбіт/с	–	20 Мбіт/с
Low 352×280	-	4 Мбіт/с	4 Мбіт/с	–	–
Кодування компонентів	4:2:0	4:2:0	4:2:0	4:2:0	4:2:0 або 4:2:2
В-кадри	Немає	Є	Є	Є	Є
Масштабованість	Немає	Немає	За С/Ш	За просторовою роздільною здатністю і за С/Ш	

Вертикальні стовпчики таблиці відповідають новій градації цифрових телевізійних систем – *профілям*. З переходом на більш високі профілі, тобто у разі просування таблицею праворуч, збільшується кількість методів кодування, з'являються нові

властивості телевізійної системи, але ускладнюються апаратура й алгоритми оброблення сигналів.

У клітинках таблиці подано максимальні значення швидкості передавання двійкових символів для варіантів стандарту, у трьох нижніх рядках таблиці наведено додаткові відомості про властивості профілів.

Розглянутий у цьому підрозділі метод кодування відноситься до основного профілю (Main у табл. 2.3). Як видно з таблиці, на основному рівні (Main), що відповідає телебаченню звичайної чіткості, швидкість передавання двійкових символів у каналі зв'язку досягає 15 Мбіт/с. Порівнявши це значення з вхідним значенням 216 Мбіт/с, що відповідає паралельному відеоструму за Рекомендацією ITU-R BT 656, бачимо, що здійснюється стиснення потоку інформації приблизно у 15 разів. Режим «Main Profile@Main Level (MP@ML)» нині широко використовують у системах цифрового телевізійного мовлення.

На більш високих рівнях головного профілю, що відповідають телебаченню високої чіткості, швидкість передавання двійкових символів у каналі зв'язку зростає до 60 або 80 Мбіт/с. Варто підкреслити, що для всіх рівнів цього профілю використовують одні й ті ж методи кодування, що свідчить про сумісність різних рівнів. На більш високих рівнях кодери та декодери повинні мати більшу швидкодію, а також більший обсяг запам'ятовувального пристрою (ЗП). Апаратура більш високих рівнів може працювати на більш низьких рівнях.

Перейдемо до розгляду інших профілів стандарту MPEG-2. Простий профіль (Simple у табл. 2.3) відрізняється від головного тим, що не містить *B*-кадрів, що спрощує апаратуру, але призводить до погіршення якості зображення за тієї ж швидкості передавання

двійкових символів. Такий профіль можна використовувати для запису зображень на магнітні або лазерні диски і для інших цілей.

Вищі профілі стандарту MPEG-2 характеризуються наявністю масштабованості. Крім того, на таких профілях можливе застосування компонентного кодування сигналів кольорового телебачення з передаванням кольорорізнцевих сигналів не лише через рядок (4:2:0), але й у кожному рядку (4:2:2), що застосовується у студійному обладнанні, зокрема для відеомонтажу.

Таким чином, стандарт MPEG-2 забезпечує параметри цілої низки цифрових телевізійних систем для різних застосувань і з різною якістю зображення, що мають у своїй основі подібні методи кодування зображення. На засадах цього стандарту можна будувати не лише телевізійні системи, але й інші системи, призначені для передавання рухомих зображень у цифровій формі (наприклад, системи телеконференцій та інтерактивного відеосервісу).

**Відмінності MPEG-1 і MPEG-2.** Стандарт MPEG-2 є більш розвиненим і розширеним порівняно зі стандартом MPEG-1, оскільки потік відеоданих MPEG-2 містить складові, яких немає в MPEG-1, наприклад масштабованість та всі пов'язані з нею особливості.

У стандарті MPEG-1 немає принципових обмежень щодо розмірів кодованих зображень і використання черезрядкової розгортки порівняно з MPEG-2. Проте MPEG-1 призначений для стиснення зображень, що рухаються, з прогресивною розгорткою, частотою кадрів до 30 Гц, кількістю рядків до 576 і кількістю елементів у рядку до 720. При цьому утворюється потік даних зі швидкістю передавання двійкових символів до 1856000 біт/с.

MPEG-1 допускає кодування телевізійних сигналів із черезрядковою розгорткою лише в кадровому режимі, а у MPEG-2 є

польовий режим, більш ефективний за наявності істотних змін зображення від першого поля кадру до другого, за якого два поля кадру кодуються незалежно. Наприклад, *I*-кадр може кодуватися як два *I*-поля або як *I*-поле і *P*-поле. Ці можливості надають додаткову гнучкість кодуванню, що допомагає досягти більш високої ефективності стиснення.

На практиці MPEG-1 використовують для стиснення зображень, що рухаються, розміром  $360 \times 240$  елементів із прогресивною розгорткою (формат SIF). Таке стиснення дає змогу записувати відеопрограми з певною втратою чіткості на компакт-диски і відтворювати їх на персональних комп'ютерах, виконуючи декодування в реальному часі виключно програмними засобами. У той же час, MPEG-2 є основою систем цифрового телевізійного мовлення, які швидко розвиваються, а також застосовується для запису кінофільмів і іншої відеоінформації на диски стандарту DVD, що забезпечують високу якість зображення та звуку.

Група MPEG починала роботу над стандартом MPEG-3, що визначає методи стиснення для телебачення високої чіткості (ТВЧ), однак до стандарту MPEG-2 було включено рівні, що відповідають ТВЧ (див. табл. 2.3), тому потреба у стандарті MPEG-3 відпала.

## **2.5. Стандарт кодування зображень MPEG-4**

Наступним після MPEG-2 проектом групи MPEG став стандарт MPEG-4 [12], опрацювання якого почали в липні 1993 року. Більшість документів стандарту MPEG-4 затверджено наприкінці 1998 – початку 1999 року. У 1999 році з'явилася друга версія MPEG-4.

Стандарт MPEG-4 охоплює такі галузі:

- цифрове телебачення та відеозв'язок;
- інтерактивну графіку, синтез зображень;

– інтерактивні мультимедійні додатки, у тому числі ті, які передаються через Інтернет.

Стандарт MPEG-4 дає можливість передавати та записувати у різні ЗП відео- і звукову інформацію з дуже великими коефіцієнтами стиснення. Спочатку його створювали для передавання рухомих зображень і звуку каналами зв'язку з вузькою частотною смугою пропускання, що потрібно як у системах відеозв'язку під час використання звичайних телефонних мереж і низькошвидкісних цифрових каналів (64 кбіт/с), так і для передавання через Інтернет. У цей час MPEG-4 набув значного поширення як спосіб записування кінофільмів і відеопрограм на лазерні диски. Крім того, новий стандарт забезпечує інтерактивність, тобто користувач має можливість керувати процесом передавання йому інформації за допомогою запитів, вибору варіантів і інших дій. Можна чекати, що можливості MPEG-4 будуть широко використовувати як у телевізійному мовленні, особливо в інтерактивному телебаченні, так і у цифрових системах прикладного телебачення.

MPEG-4 – складна система, повний опис якої вимагає окремої книги, тому наведемо короткий огляд його основних понять.

**Об'єкти й сцени.** Найважливішою особливістю MPEG-4 є *об'єктно-орієнтований* підхід, суть якого полягає в тому, що зображення зі звуковим супроводженням подається як сукупність відео- та аудіооб'єктів.

*Відеооб'єктами* (VO – visual object) можуть бути зображення людей і предметів, що переміщуються перед нерухомим фоном, і сам нерухомий фон. Звичайне телевізійне зображення може бути єдиним відеооб'єктом.

*Аудіооб'єктами* (АО - audio object) можуть бути голоси людей, музика, інші звуки. Пов'язані відео- та аудіооб'єкти, наприклад

зображення людини і його голос, утворюють аудіо-візуальний об'єкт (AVO – audio-visual object). Відео- та аудіооб'єкти становлять *сцену*. MPEG-4 містить спеціальну мову для опису сцен – BIFS (Binary Format for Scenes – двійковий формат для сцен).

Опис сцени має ієрархічну структуру, верхнім рівнем якої є сцена в цілому, що зазвичай має нерухомий фон (наприклад, зображеннями землі, трави, дерев тощо). У сцені можуть бути різні аудіовізуальні об'єкти (декілька людей), кожний з яких містить відеооб'єкт – рухоме зображення людини, а також аудіооб'єкт – голос цього персонажа.

Опис кожної сцени включає дані про координати об'єктів у просторі й про їх прив'язку до часу. Відеооб'єкти можуть розміщуватися в різних планах сцени та перекривати один одного. Сцена може містити всі об'єкти, інформація про які надходить у прийнятому потоці даних, або лише частину цих об'єктів. В інтерактивному режимі користувач може впливати на розвиток сцени, подаючи відповідні команди.

MPEG-4 дає також можливість передавати користувачеві додаткову інформацію про об'єкти, наприклад у вигляді вікна з текстом, що з'являється, коли користувач обрав який-небудь об'єкт у сцені.

**Кодування відеооб'єктів.** Відеооб'єкти, що утворюють сцену, кодуються окремо. Кодовані відеодані всіх відеооб'єктів поєднуються в послідовність візуальних об'єктів (VOS – Visual Object Sequence). Відеооб'єкт може мати кілька шарів або рівнів, кожному з яких відповідає елементарний потік відеоданих. Наявність декількох шарів забезпечує можливість масштабованості. Базовий шар дає зображення відеооб'єкта з базовою якістю, додаткові або поліпшувальні (enhanced) шари допомагають отримувати зображення

з кращою просторовою роздільною здатністю, меншими спотвореннями через стиснення тощо.

Зображення відеооб'єкта, яке змінюється в часі, передається у вигляді послідовності його нерухомих зображень у дискретні моменти часу. Ці нерухомі зображення називають *площинами відеооб'єкта* (VOP – Visual Object Plane). Якщо відеооб'єктом є звичайне телевізійне зображення, то площинами цього відеооб'єкта можуть бути звичайні кадри. VOP поєднують у групи (GOV). VOP, як і кадри в MPEG-2, можуть кодуватися незалежно від інших VOP (*I-VOP*) або з передбаченням та компенсацією руху (*P-VOP* і *B-VOP*).

На відміну від MPEG-1, MPEG-2, які мають фіксований алгоритм кодування, у MPEG-4 використовується низка методів кодування, які містять як алгоритми, подібні до застосовуваних у MPEG-1, MPEG-2, так і принципово нові методи кодування, ґрунтуються на понятті «відеооб'єкт». Вибір того або іншого методу кодування в конкретному випадку визначається типом відеооб'єкта й заданим коефіцієнтом стиснення інформації. MPEG-4 дає змогу ефективно виконувати стиснення як натуральних, так і синтетичних зображень, та поєднувати їх під час відтворення.

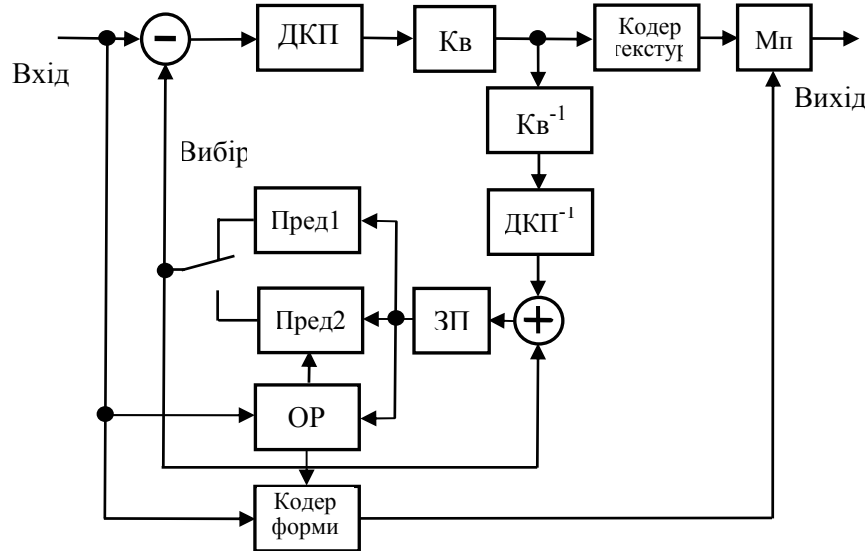
Узагальнену структурну схему відеокодера MPEG-4 для натуральних зображень наведено на рис. 2.10.

На схемі позначено: ДКП – блок прямого дискретного косинусного перетворення; ДКП<sup>-1</sup> – блок оберненого дискретного косинусного перетворення; Кв - квантувач; Кв<sup>-1</sup> – деквантувач, тобто блок, що виконує обернену квантуванню операцію; ЗП – запам'ятовувальний пристрій; Пред1, Пред2 – блоки, що формують передбачені зображення у різних режимах кодування; ОР – блок оцінки руху й формування векторів руху; Мп – мультиплексер. Крім того, на схемі показано блоки «Кодер форми» і «Кодер текстур»,



пристрої додавання та віднімання, а також перемикач «Вибір», за допомогою якого здійснюється підключення одного з блоків передбачення залежно від методу кодування.

На вхід кодера надходять відеодані, наприклад цифровий телевізійний сигнал, на виході формується елементарний потік відеоданих.



**Рис. 2.10.** Структурна схема відеокодера MPEG-4

Коротко розглянемо основні методи кодування натуральних зображень.

1. Відеооб'єкти, що являють собою прямокутні зображення (наприклад, звичайні телевізійні кадри), кодуються методом, аналогічним застосовуваному в MPEG-1, MPEG-2, тобто з використанням гібридного кодування (див. підрозд. 2.4). Метод включає передбачення з оцінкою та компенсацією руху для

макроблоків  $16 \times 16$  пікселів і ДКП похибки передбачення у блоках  $8 \times 8$  пікселів. Будемо вважати, що передбачення в цьому разі виконується у блоці Пред1.

Цей вид кодування має два рівні за швидкістю передавання двійкових символів у вихідному потоці даних. Рівень дуже низької швидкості передавання VLBV (Very Low Bitrate Video) призначено для передавання зображень з низькою просторовою роздільною здатністю (формати QCIF і SQCIF) і зниженою частотою кадрів (10...15 Гц) каналами зв'язку з вузькою частотною смугою пропускання (швидкість передавання двійкових символів 5...64 кбіт/с). Цей рівень можна використовувати у відеотелефонному зв'язку з невисокою якістю зображення.

Рівень високої швидкості передавання (High bitrate) призначено для передавання зображень з більш високою просторовою роздільною здатністю, аж до формату за Рекомендацією 601, різними каналами зв'язку зі швидкостями передавання двійкових символів від 64 кбіт/с до 10 Мбіт/с. Цей рівень можна використовувати у відеозв'язку з високою якістю зображення і для передавання телевізійних програм.

2. Кодування, в основі якого лежить зміст (content-based coding), за допомогою якого досягається набагато більше стиснення зображень завдяки врахуванню властивостей відеооб'єктів, які є у сцені.

Однією з можливостей цих методів є кодування відеооб'єктів складної форми. Наприклад, як відеооб'єкт може бути використана ділянка зображення, що відрізняється яскравістю або кольором та може пересуватися й деформуватися. Під час формування передбаченого зображення з компенсацією руху зміщуються не прямокутні макроблоки, а виділені ділянки, які, до того ж, можуть

змінювати свою форму. При цьому похибка передбачення є значно меншою та обсяг інформації, що міститься у різниці передбаченого й дійсного зображень чергового кадру, істотно зменшується. У кодері, поданому на рис. 2.10, такий варіант передбачення виконується у блоці Пред2.

Форму відеооб'єкта подають у вигляді матриці. У разі бінарного кодування елементи матриці, що відповідають елементам зображення відеооб'єкта, прирівнюються до одиниці, а відповідні елементи зображення, що перебувають поза відеооб'єктом, – до нуля. У разі градаційного кодування елементи матриці набувають більше значень, що дає змогу описувати такі властивості об'єкта, як прозорість. Елементи зазначеної матриці визначаються й кодуються в «кодері форми», після чого вони через мультиплексер Мп надходять на вихід (див. рис. 2.10).

Дані про форму відеооб'єкта містять значно більше двійкових символів, ніж простий вектор руху. Проте загальний вииграш щодо зменшення обсягу переданої інформації, якщо порівнювати з MPEG-1, MPEG-2, є суттєвим.

3. Для стиснення зображень нерухомого фону й текстур об'єктів великих розмірів використовують метод кодування, в основі якого лежить вейвлет-перетворення, яке розглянемо у підрозд. 2.7. Цей метод забезпечує високий ступінь стиснення й багатоступеневу масштабованість за просторовою роздільною здатністю.

4. Зображення заднього плану, які не змінюються або майже не змінюються, можуть передаватися як *спрайт* (sprite), повне зображення якого передається один раз. Далі передаються лише параметри глобального руху, що визначають переміщення телевізійної камери відносно заднього плану.

Синтетичні відеооб'єкти, створені за допомогою засобів

машинної графіки, можуть кодуватися зазначеними методами для натуральних зображень, однак значно ефективніше використати їх параметричний опис.

У стандарті MPEG-4 використовується модель обличчя людини, побудована на основі сітки з трикутних комірок, які заповнюються текстурою. Є також тривимірна модель тіла людини у вигляді тривимірної сітки. Двовимірні зображення людини можна отримати, побудувавши проекцію тривимірної моделі на задану площину.

Форма, текстура й вираз обличчя у статичі визначаються параметрами FDP (Facial Definition Parameters), а в динаміці – параметрами FAP (Facial Animation Parameters). Для тіла у статичі задають параметри BDP (Body Definition Parameters), а у динаміці – BAP (Body Animation Parameters). Статичні параметри FDP і BDP передаються на початку сеансу зв'язку. Для відтворення міміки обличчя й рухів тіла співрозмовника у процесі розмови передаються динамічні параметри FAP і BAP.

Синтетичні зображення обличчя й тіла людини можна використовувати у системах відеозв'язку замість дійсних зображень співрозмовника. Використання моделі вимагає істотно меншої швидкості передавання двійкових символів, ніж кодування реального зображення.

У деяких випадках можна у приймальній частині системи отримати інформацію про зміни зображення об'єкта на основі іншої інформації. Так, наприклад, рухи рота і міміка обличчя людини, що говорить, зумовлені вимовленими словами і можуть бути синтезовані на основі прийнятого голосу співрозмовника. При цьому потрібна для здійснення відеозв'язку швидкість передавання двійкових символів стає ще меншою.

Крім обличчя і тіла, можуть синтезуватися довільні двовимірні

зображення також у вигляді сіток з трикутними комірками, які заповнені текстурою.

Стандартом MPEG-4 передбачено багаторівневу масштабованість за просторовою роздільною здатністю, за часом і якістю зображення, а також засоби, що забезпечують працездатність системи передавання відеоінформації за наявності перешкод у каналі зв'язку.

**Декодування і відтворення.** Структурну схему декодувальної частини системи за стандартом MPEG-4 наведено на рис. 2.11, де показано демультимплексер ДМп, буферні запам'ятовувальні пристрої БЗП1 і БЗП2, декодери ДКд і блок об'єднання БОб.

На вхід надходить транспортний потік даних, з якого в демультимплексері виділяються елементарні потоки, дані кожного з них записуються у відповідний БЗП1. Призначення БЗП1 – накопичувати дані, які нерівномірно надходять з каналів зв'язку, та передавати їх на декодер.

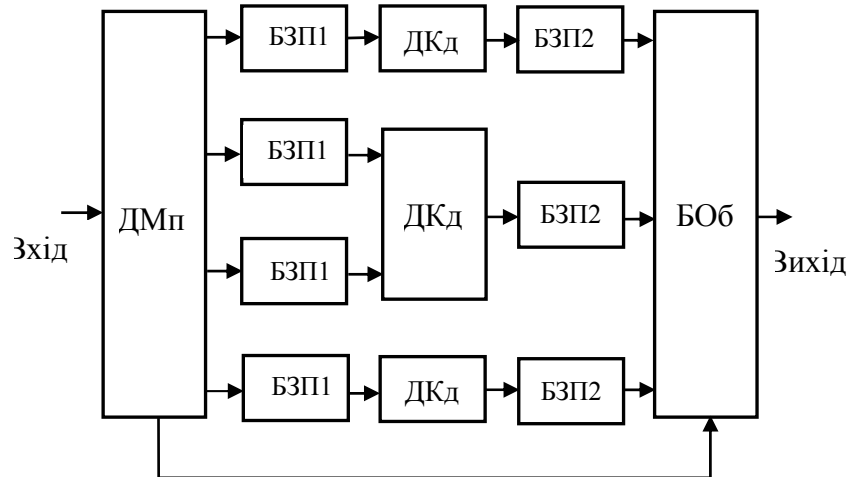


Рис. 2.11. Декодувальна частина MPEG-4

Далі виконується декодування елементарних потоків, після чого дані відео- та аудіооб'єктів записуються у БЗП2. Елементарні потоки, що належать одному об'єкту, можуть декодуватися разом. У разі демультимплексування з загального потоку даних виділяються також опис сцени й мітки часу, що надходять на блок об'єднання.

Дані окремих об'єктів зчитуються з відповідних БЗП2 і з них у блоці об'єднання формуються цифрові сигнали зображення і звуку, що надходять далі на відтворювальні пристрої (на рис. 2.11 не показано). При цьому забезпечується синхронізація всіх відео- і аудіооб'єктів.

**Типи об'єктів, профілі й рівні MPEG-4.** Поняття «рівень» й «профіль» було введено ще у стандарті MPEG-2, у той час як поняттям «тип об'єкта» оперує лише MPEG-4, яким передбачено такі типи відеооб'єктів:

- простий – прямокутний відеооб'єкт, під час кодування якого використовуються VOP *I*- та *P*-типів;
- простий масштабований, у якого можливе масштабування як за просторовою роздільною здатністю, так і за часовою;
- базовий, у якого порівняно з простим можливе використання VOP *B*-типу;
- основний, у якого порівняно з базовим можливі градаційне кодування форми, спрайти та черезрядкова розгортка;
- *N*-бітовий, у якого порівняно з базовим є можливість задавати кількість двійкових розрядів у межах 4...12;
- нерухома масштабована текстура, яка кодується з використанням вейвлет-перетворення;
- двовимірна сітка з можливістю анімації, за допомогою якої може подаватися натуральне зображення з подальшим його змінюванням завдяки переміщенню вузлів сітки;

- базова текстура з можливістю анімації нерухомих зображень;
- обличчя людини із засобами анімації.

Профілі у MPEG-4 розрізняють за складом об'єктів, за можливою кількістю відеооб'єктів у сцені та за максимальною швидкістю передавання двійкових символів у потоці даних на виході. Різні профілі застосовують у різних галузях. Рівні, як і в MPEG-2, визначаються кількістю елементів зображення у сцені. Параметри трьох із дев'яти профілів, передбачених першою версією MPEG-4, наведено у табл. 2.4. Для звукової інформації в першій версії MPEG-4 передбачено 15 типів об'єктів, що розрізняються використанням методом кодування і наявністю масштабованості.

У другій версії MPEG-4 введено додаткові профілі як для зображень, так і для звуків, а також графічні профілі та профілі опису сцени та об'єктів.

**Таблиця 2.4.** Профілі стандарту MPEG-4

Профіль	Рівень	Розмір сцени	Швидкість передавання	Типи об'єктів	Максимальна кількість об'єктів
Простий	L1	QCIF	64 кбіт/с	Простий	4
	L2	CIF	128 кбіт/с		4
	L3	CIF	384 кбіт/с		4
Базовий	L1	QCIF	384 кбіт/с	Простий, базовий	4
	L2	CIF	2 Мбіт/с		16
Основний	L2	CIF	2 Мбіт/с	Простий, базовий, основний	16
	L3	Реком. 601	15 Мбіт/с		32
	L4	1920×1088	38,4 Мбіт/с		32

## 2.6. Інші стандарти стиснення зображень

Розглянемо деякі стандарти кодування відеоінформації, які визначені у Рекомендаціях ITU-T H.261, H.262 та H.263 і застосовуються у системах відеозв'язку, комп'ютерному обробленні зображень і інших галузях [13–15].

Рекомендація ITU-T H.261 визначає методи кодування і декодування відеосигналів для передавання каналами зв'язку з порівняно вузькою частотною смугою пропускання зі швидкостями передавання двійкових символів, рівними  $n \cdot 64$  кбіт/с, де  $n = 1 \dots 30$ . Основна галузь застосування Рекомендації H.261 – системи комп'ютерного відеозв'язку.

Зображення, які кодуються, можуть мати формати CIF (352×288) або QCIF (176×144). Формат дискретизації 4:2:0, тобто кількість елементів кольорорізнцевих сигналів як по вертикалі, так і по горизонталі, у два рази менша, ніж кількість елементів сигналу яскравості. Розгортка прогресивна, з частотою кадрів 29,97 Гц, причому можливе пропускання 1, 2 або 3 кадри між кожними двома переданими кадрами, що допомагає знизити частоту кадрів до 15, 10 і 7,5 Гц відповідно. Квантування сигналу яскравості та кольорорізнцевих сигналів відбувається відповідно до Рекомендації 601.

Методи стиснення відеоінформації, які використовуються у H.261, багато в чому схожі з застосовуваними у MPEG-1 і MPEG-2. Основною одиницею відеоінформації, що кодується, є макроблок розміром 16×16 пікселів, що містить 4 блоки 8×8 елементів сигналу яскравості і по одному блоку 8×8 елементів кольорорізнцевих сигналів. 33 макроблоки становлять групу макроблоків. Кадр формату CIF містить 12 груп, а кадр формату QCIF – 3 групи макроблоків.



У кожному макроблоці може відбуватися кодування всередині кадру або міжкадровим методом. Рекомендація H.261 не містить правил вибору методу кодування для макроблоків, залишаючи це питання на розсуд розробників. Таким чином, можна довільно задавати кількість і положення кадрів, які цілком передаються з кодуванням усередині кадру, а для інших кадрів обирати найкращий метод кодування макроблоків з різними властивостями.

Кодування макроблока методом усередині кадру, як і у JPEG, MPEG-1 і MPEG-2, включає операції поблокового ДКП, квантування коефіцієнтів ДКП, перетворення матриці коефіцієнтів ДКП у їх послідовність за допомогою зчитування в зигзагоподібному порядку, кодування отриманої послідовності парами чисел (run-length coding) і кодування кодами зі змінною довжиною кодових слів (кодування за Хаффманом). Ступінь стиснення визначається одним з 32 можливих значень кроку квантування, причому для коефіцієнтів  $C(0, 0)$ , що відповідають постійним складовим сигналів, крок квантування фіксований.

Міжкадрове кодування полягає у передбаченні кодованого макроблока за рахунок пошуку відповідної ділянки в попередньому кадрі, отримання похибки передбачення (різниці дійсного й передбаченого макроблоків) і кодування похибки передбачення методом, який використовується у разі кодування всередині кадру, але з іншими параметрами квантування. Вектор руху, що показує зміщення відповідної ділянки в попередньому кадрі, за кожною координатою визначається з точністю до одного пікселя в діапазоні –15...15 пікселів. Вектори руху передаються в потоці даних. Відзначимо, що передбачення здійснюється лише за попереднім кадром, а *B*-кадрів, наявних у MPEG-1, MPEG-2, тут немає.

Структура потоку відеоданих на виході кодера містить заголовки

кадрів, груп макроблоків, макроблоків і дані, які передаються поблоково у кожному макроблоці. Об'єднання кадрів у групи не передбачено. У декодері потік даних розділяється на кадри, групи макроблоків і макроблоки, кожний з яких декодується аналогічно до стандарту MPEG-2.

Щоб оцінити можливості передавання зображень із використанням Рекомендації H.261, розглянемо приклад. Нехай нам потрібно передавати зображення формату QCIF із частотою кадрів 10 Гц. Такі параметри прийнятні для простого відеотелефонного зв'язку. Швидкість передавання двійкових символів до стиснення становить 2970 кбіт/с. Для передавання каналом зв'язку з пропускнуою здатністю 64 кбіт/с має відбуватися стиснення у 46 разів, але таке стиснення призведе до істотних спотворень зображення. Якщо ж пропускна здатність каналу зв'язку 128 кбіт/с, то необхідне стиснення у 23 рази, що цілком припустимо.

Рекомендація H.263 більш розвинена порівняно з Рекомендацією H.261 та має такі основні відмінності:

- крім зображень форматів CIF і QCIF, можуть кодуватися і передаватися зображення форматів SQCIF, 4CIF і 16CIF;
- вектори руху визначаються з точністю до половини пікселя, і обмежень щодо розмірів зсувів немає;
- допускається використання *B*-кадрів, під час кодування яких для передбачення залучають наступні за ними кадри;
- використовується більш досконалий метод кодування зі змінною довжиною кодових слів (арифметичне кодування).

Рекомендація H.262 є суттєво більш змістовною, ніж H.261 і H.263, та багато в чому збігається зі стандартом MPEG-2. Відповідно до неї рухомі зображення, що кодуються, можуть мати як

прогресивну, так і через рядкову розгортку, кількість пікселів від  $352 \times 288$  до  $1920 \times 1152$ , частоту кадрів 30 Гц, а на верхніх рівнях 60 Гц. Передбачено рівні і профілі, а також можлива масштабованість декількох типів. Рекомендацію H.262 можна використовувати не лише для систем відеозв'язку, але і для передавання програм телевізійного мовлення різними каналами зв'язку та у цифрових системах прикладного телебачення.

## **2.7. Стиснення зображень із застосуванням вейвлет-перетворення**

Перспективний підхід до стиснення зображень ґрунтується на розкладенні зображення на складові за роздільною здатністю. Як відомо, зображення складається з великих деталей, що становлять основу його змісту, і дрібної структури. У багатьох випадках можна розділити зазначені складові, щоб передавати їх окремо, а у приймальній частині телевізійної системи об'єднати ці складові в повне зображення.

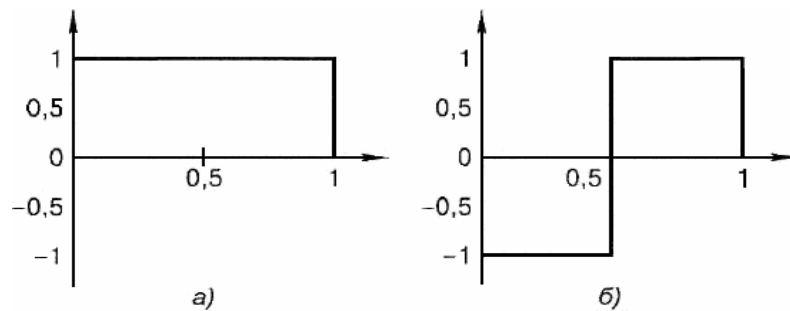
Останнім часом все більшої популярності набуває новий підхід до стиснення зображень, побудований на застосуванні біортогонального вейвлет-перетворення (Wavelet-Transform – WT). Термін «wavelet» дослівно перекладають як «маленька хвиля» [16].

На відміну від гармонічних ортогональних функцій, які визначені на всьому інтервалі зміни аргументу, базисні функції вейвлет-перетворення відмінні від нуля лише на обмеженому відрізку цього інтервалу. Повний набір таких функцій, що становлять базис перетворення, покриває весь інтервал зміни аргументу.

Результатом вейвлет-перетворення є набір коефіцієнтів, відновлення сигналу за якими називається оберненим вейвлет-перетворенням (Inverse Wavelet Transform – IWT).

Простим прикладом вейвлет-перетворення є розкладання сигналу за базисом Хаара (Haar). Базисну функцію перетворення Хаара, зображену на рис. 2.12, *a*, називають scale-function або scaling-function. Ортогональний базис Хаара отримують за допомогою зсуву базисної функції з кроком, рівним одиниці, в усі можливі положення на осі абсцис.

Функцію, наведену на рис. 2.12, *б*, називають вейвлетом Хаара (Haar Wavelet). Вейвлети Хаара, які зсуваються на осі абсцис, також утворюють ортогональний базис для простору сигналів *деталей*. Інші варіанти вейвлет-перетворення також мають базисні функції (scaling-functions) і вейвлети, але більш складні за формою.



**Рис. 2.12.** Базисна функція Хаара (*a*) і вейвлет Хаара (*б*)

Розглянемо спочатку одновимірний випадок. Нехай є вхідний цифровий сигнал  $x(n)$ ,  $n = 0, 1, \dots, N-1$ . У результаті перетворення Хаара утворюються два цифрових сигнали  $x_1(m)$  і  $x_2(m)$ ,  $m = 0, 1, \dots, N/2-1$  ( $N$  передбачається парним), відповідно до співвідношень

$$\begin{aligned} x_1(m) &= [x(2m+1) + x(2m)]/2; \\ x_2(m) &= x(2m) - x(2m+1). \end{aligned} \quad (2.20)$$

Відліки сигналу  $x_1(m)$  отримують через усереднення пар суміжних відліків вихідного сигналу  $x(n)$ , тому сигнал  $x_1(m)$  містить

інформацію зі зменшеною у 2 рази роздільною здатністю.

У той же час, сигнал  $x_2(m)$ , відліки якого дорівнюють різницям значень суміжних відліків вихідного сигналу, містить інформацію про дрібні деталі вихідного сигналу. Саме для сигналу  $x_2(m)$  ортогональним базисом є вейвлети Хаара. Кожен із сигналів  $x_1(m)$  і  $x_2(m)$  містить  $N/2$  відліків, і разом вони несуть усю інформацію, яка міститься у вихідному сигналі.

Обернене перетворення Хаара визначається співвідношеннями, що впливають зі співвідношення (2.20), де перше дає змогу відновити непарні відліки вихідного сигналу, а друге – парні:

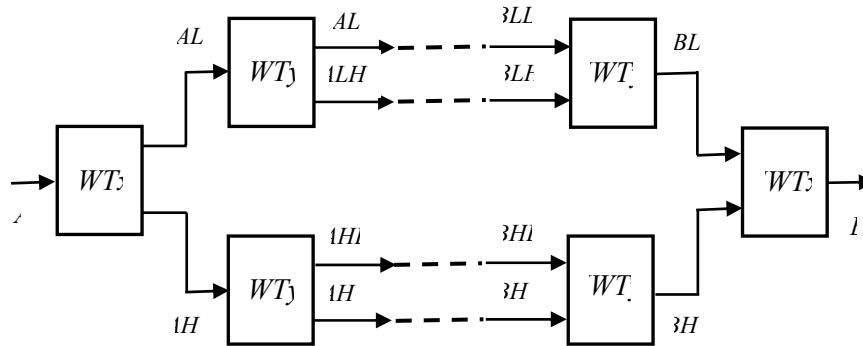
$$\begin{aligned}x_1(m+1) &= [2x_1(m) - x_2(m)]/2; \\ x(2m) &= [2x_1(m) - x_2(m)]/2.\end{aligned}\tag{2.21}$$

Далі перетворення Хаара можна застосувати до сигналу  $x_1(m)$ , який при цьому буде розкладений на сигнал  $x_{11}(k)$ , що відповідає ще більш низькій роздільній здатності, і сигнал  $x_{12}(k)$ , що містить деталі, потрібні для відновлення сигналу  $x_1(m)$ . Сигнали  $x_{11}(k)$  і  $x_{12}(k)$  містять по  $N/4$  відліків. Потім перетворення Хаара можна застосувати до сигналу  $x_{11}(k)$  і т. д. Відповідними оберненими перетвореннями можуть бути послідовно відновлені сигнали з більш високою роздільною здатністю, аж до вхідного сигналу.

Слід зазначити, що у процесі прямого перетворення Хаара (як і прямого вейвлет-перетворення іншого типу) кількість відліків у сигналах, які утворюються на кожній стадії перетворення, зменшується в 2 рази, а кількість самих сигналів зростає. При цьому сумарна кількість відліків у всіх отриманих сигналах завжди дорівнює  $N$ . Під час оберненого перетворення кількість відліків у сигналах, що відновлюються на кожній стадії, збільшується у 2 рази.

Вейвлет-перетворення зображення здійснюється послідовно за двома координатами (рис. 2.13). Вхідне зображення визначається

матрицею  $A$ , що містить  $N^2$  елементів. У блоках  $WT_x$  та  $WT_y$  виконуються вейвлет-перетворення одновимірних цифрових сигналів, наприклад перетворення Хаара, у блоках  $IWT_x$  і  $IWT_y$  – відповідні обернені перетворення.



**Рис. 2.13.** Аналіз і синтез зображення з використанням вейвлет-перетворення

На першому етапі відбувається вейвлет-перетворення за горизонтальною координатою  $x$ , при цьому елементи зображення  $A$ , розташовані в кожному рядку матриці, послідовно надходять у блок  $WT_x$ . У результаті формуються два зображення  $AL$  і  $AH$ , кожне з яких являє собою матрицю, що містить  $N/2$  елементів по горизонталі й  $N$  елементів по вертикалі. Сумарна кількість елементів у цих зображеннях залишилася рівною  $N^2$ . У зображенні  $AL$  роздільна здатність по горизонталі зменшена у два рази. Зображення  $AH$  містить інформацію про деталі, що мають малі розміри за координатою  $x$ .

Потім кожне із зображень  $AL$  і  $AH$  піддається вейвлет-перетворенню за вертикальною координатою  $y$ , при цьому на блоки  $WT_y$  послідовно надходять елементи зображення, розташовані у

стовпцях відповідних матриць. У результаті утворюються чотири матриці, кожна з яких містить  $(N/2)^2$  елементів. Сумарна кількість елементів, як і раніше, дорівнює  $N^2$ .

Матриця  $ALL$  відповідає зображенню, подібному до вхідного, але зі зменшеними у два рази роздільними здатностями за обома координатами. Матриця  $ALH$  містить деталі вхідного зображення, що мають порівняно великі розміри по горизонталі, але мінімальні по вертикалі. Матриця  $AHL$  містить деталі зображення, які мають порівняно великі розміри по вертикалі, але мінімальні по горизонталі. Нарешті, матриця  $AHH$  містить деталі зображення, що мають малі розміри за обома координатами.

Розкладення на діапазони за роздільною здатністю дає змогу аналізувати або передавати зображення з різними ступенями деталізації. Для отримання найбільш грубого зображення достатньо використати лише інформацію, що міститься в матриці  $ALL$ . Додавання інших діапазонів збільшує детальність зображення. Як і для одновимірних сигналів, кількість ступенів двовимірного вейвлет-перетворення може бути збільшена. Матриця  $ALL$  може бути перетворена в чотири матриці, кожна з яких містить  $(N/4)^2$  елементів. Та з цих матриць, що містить зображення зі зменшеними в чотири рази просторовими роздільними здатностями за обома координатами, може бути перетворена у чотири матриці розміром  $(N/8)^2$  елементів і т. д.

Після розкладення вхідного зображення  $A$  на діапазони за роздільною здатністю може виконуватися оброблення кожного з цих діапазонів, у результаті чого матриця  $ALL$  перетворюється у  $BLL$  і т. д. Обернені вейвлет-перетворення виконуються у зворотному порядку (рис. 2.13), у результаті формується зображення  $B$  розміром  $N \times N$  елементів.

Вейвлет-перетворення є ефективним засобом стиснення зображень. Зазвичай для подання даних, що містяться в матрицях  $ALH$ ,  $AHL$  і  $ANH$ , може використовуватися менша кількість двійкових розрядів, тобто більш грубе квантування, ніж для вхідного зображення. Це зумовлено тим, що сигнали від деталей з малими розмірами мають менший розмах. Дані з усіх матриць після квантування поєднуються в загальний потік даних. У приймальній частині відновлюється початкова кількість двійкових розрядів.

Вейвлет-перетворення забезпечує більш високий коефіцієнт стиснення нерухомих зображень зі збереженням їх прийнятної якості, ніж поширений стандарт JPEG, однак за умови, що використовується не перетворення Хаара, а інші типи вейвлет-перетворень з більш складними за формою базисними функціями та вейвлетами.

Важливою перевагою вейвлет-перетворення перед ДПФ і ДКП є роздільне виконання перетворень за двома координатами. Загальна кількість обчислювальних операцій під час вейвлет-перетворення зображення, яке містить  $N \times N$  елементів, пропорційна  $N^2$ , тобто набагато менше, ніж для ДПФ і ДКП, навіть у разі застосування для їх реалізації швидких алгоритмів.

### **Контрольні запитання**

- 1. Як розраховують ширину смуги частот цифрового телевізійного сигналу в послідовній формі та потрібну для його передавання ширину смуги частот каналу зв'язку?*
- 2. Обґрунтуйте необхідність стиснення телевізійного зображення для його передавання стандартними каналами зв'язку.*
- 3. Що таке психофізіологічна надмірність телевізійного*



зображення?

4. Чому ДПФ містить скінченну кількість частотних складових?

5. Чому у багатьох випадках зручніше застосовувати ДКП, ніж ДПФ?

6. У якій послідовності кодуються за стандартом JPEG блоки кольорового зображення?

7. Як кодуються коефіцієнти ДКП за стандартом JPEG?

8. Чому квантування коефіцієнтів ДКП призводить до менш помітних спотворень, ніж квантування самого зображення?

9. Яким чином у стандарті JPEG здійснюється керування ступенем стиснення?

10. Поясніть сутність кодування зі змінною довжиною кодових слів.

11. Що означає термін «гібридне кодування» за стандартами MPEG-1, MPEG-2?

12. Навіщо перед кодуванням згідно зі стандартами MPEG-1, MPEG-2 відбувається переставлення кадрів в GOP?

13. Чим розрізняються кадровий і польовий режими кодування за стандартами MPEG-1, MPEG-2?

14. Чому для В-кадрів досягається найбільший ступінь стиснення?

15. Яке призначення буферного ЗП у кодері MPEG-2?

16. Що таке масштабованість стандарту MPEG-2?

17. Що таке рівні та профілі стандарту MPEG-2?

18. Що таке відео- та аудіооб'єкти за стандартом MPEG-4?

19. Які методи використовуються за стандартом MPEG-4 для передавання зображень облич людей?

20. Опишіть процес декодування потоку даних MPEG-4.

21. Назвіть типи відеооб'єктів, передбачені у стандарті MPEG-4.
22. Що таке профілі стандарту MPEG-4?
23. Де можуть застосовуватися методи кодування і декодування телевізійних сигналів згідно з Рекомендаціями H.261, H.262 і H.263?
24. Що відбувається із зображенням під час здійснення вейвлет-перетворення?
25. За рахунок чого може бути досягнуто стиснення зображення у процесі вейвлет-перетворення?
26. Назвіть переваги вейвлет-перетворення перед стандартом JPEG та перед ДПФ і ДКП.

### 3. ЦИФРОВЕ ОБРОБЛЕННЯ ТЕЛЕВІЗІЙНИХ СИГНАЛІВ

#### 3.1. Цифрова фільтрація одновимірних сигналів

Цифрова фільтрація – один з найважливіших методів оброблення одновимірних і багатовимірних цифрових сигналів. За допомогою цифрових фільтрів у телебаченні вирішують завдання щодо зменшення впливу шумів і перешкод, підвищення суб'єктивної якості зображення тощо. При цьому використовують як одновимірну, так і двовимірну (просторову) фільтрацію.

Лінійний цифровий фільтр перетворює послідовність відліків  $x(i)$  вхідного сигналу в послідовність відліків  $y(i)$  сигналу на виході відповідно до правила

$$y(i) = a_0 x(i) + a_1 x(i-1) + \dots + a_m x(i-m) + b_1 y(i-1) + \dots + b_n y(i-n). \quad (3.1)$$

Значення  $m$  і  $n$  можуть бути будь-якими натуральними числами, найбільше з яких називають порядком цифрового фільтра.

Коефіцієнти  $a_j$ ,  $b_j$  можуть бути будь-якими дійсними числами, як додатними, так і від'ємними. Частина цих коефіцієнтів може дорівнювати нулю. Властивості цифрового фільтра повністю визначаються його коефіцієнтами.

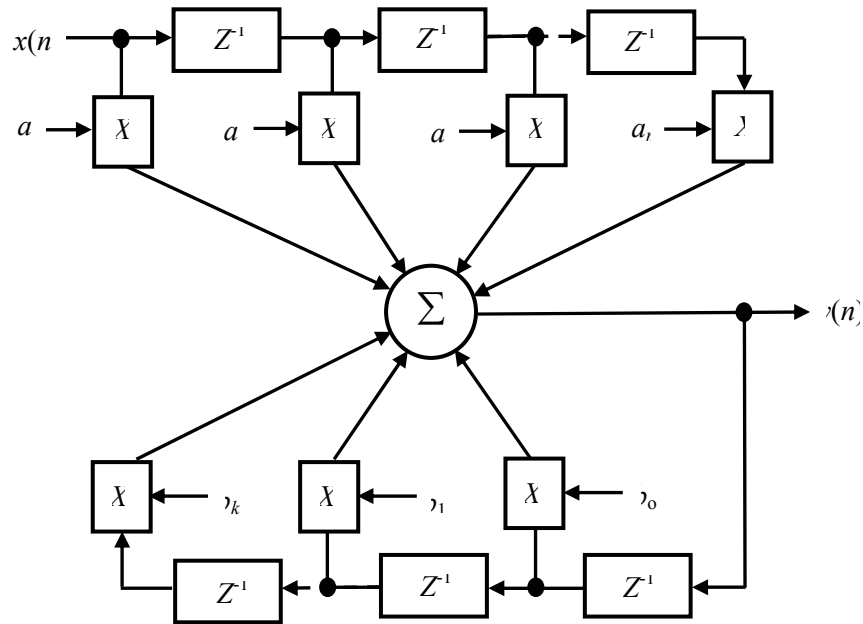
Структурну схему цифрового фільтра, що виконує перетворення (3.1), подано на рис. 3.1. Відзначимо, що це один з можливих варіантів структури цифрового фільтра з таким набором коефіцієнтів. Кожний блок  $Z^{-1}$  здійснює затримку сигналу на один відлік. Таке позначення блоку пов'язано із  $z$ -перетворенням, яке є зручним засобом опису та аналізу цифрових фільтрів. Опускаючи математичні викладки, з якими можна ознайомитися у праці [7], відразу запишемо результат  $z$ -перетворення співвідношення (3.1):

$$Y(z) = (a_0 + a_1 z^{-1} + \dots + a_m z^{-m})X(z) + (b_1 z^{-1} + \dots + b_n z^{-n})Y(z), \quad (3.2)$$

де  $X(z)$  і  $Y(z)$  –  $z$ -перетворення сигналів  $x(i)$  і  $y(i)$ . Затримку цифрового сигналу на  $k$  відліків у результаті  $z$ -перетворення замінюють множенням на  $z^{-k}$ .

Знайдемо з рівняння (3.2)  $Y(z)$

$$Y(z) = \frac{a_0 + a_1 z^{-1} + \dots + a_m z^{-m}}{1 - b_1 z^{-1} - \dots - b_n z^{-n}} X(z) = H(z)X(z). \quad (3.3)$$



**Рис. 3.1.** Структурна схема цифрового фільтра для одновимірних сигналів

Функцію  $H(z)$  називають передавальною або системною функцією цифрового фільтра. Якщо у виразі (3.3) замінити всі  $z^{-k}$

на  $e^{-jk2\pi f}$ , то утвориться комплексна частотна передавальна функція фільтра  $H(j2\pi f)$ , модуль якої є амплітудно-частотною характеристикою (АЧХ) фільтра  $H(2\pi f)$ , де  $f$  – безрозмірна частота, яка дорівнює відношенню частоти сигналу до частоти дискретизації і набуває значень в інтервалі  $0 \dots 0,5$  відповідно до умови Котельникова. Як і в аналоговій техніці, АЧХ визначає залежність відношення амплітуди сигналу на виході до амплітуди вхідного сигналу від частоти, причому вхідний сигнал має дискретний синусоїдальний вигляд.

Іншою важливою характеристикою цифрового фільтра є імпульсна характеристика  $h(n)$  – реакція фільтра на вхідний сигнал, у якому один відлік дорівнює одиниці, а всі інші – нулю. Операцію обчислення відліків вихідного сигналу за формулою (3.1) називають згорткою вхідного сигналу з імпульсною характеристикою фільтра.

Якщо відомо всі коефіцієнти цифрового фільтра у співвідношенні (3.1), то й імпульсну характеристику, і передавальну функцію, і АЧХ фільтра можна розрахувати. Обернена задача – синтез цифрового фільтра, тобто обчислення коефіцієнтів цифрового фільтра, що має задану АЧХ, – значно складніша. Ефективним інструментом синтезу цифрових фільтрів може бути програма MatLAB [17].

Якщо у формулі (3.1) всі коефіцієнти  $b_j = 0$ , то цифровий фільтр називають нерекурсивним, передавальна функція якого має вигляд

$$H(z) = a_0 + a_1 z^{-1} + \dots + a_m z^{-m}. \quad (3.4)$$

Нехай  $x(0) = 1$ , а інші –  $x(i) = 1$ , тоді за співвідношенням (3.1) отримуємо  $y(0) = a_0$ ,  $y(1) = a_1$ , ...,  $y(m) = a_m$ , а решта відліків сигналу на виході дорівнює нулю. Отже, імпульсною характеристикою нерекурсивного фільтра є послідовність його

коефіцієнтів. Реакція такого фільтра на одиничний імпульс займає  $m + 1$  відлік вихідного сигналу, тому нерекурсивні фільтри ще називають фільтрами з кінцевою імпульсною характеристикою (КІХ-фільтрами) або трансверсальними фільтрами.

Як ілюстрацію можливостей нерекурсивних цифрових фільтрів на рис. 3.2 подано задану АЧХ трапецеїдальної форми та АЧХ двох синтезованих нерекурсивних ФНЧ. На графіках по горизонталі відкладено частоту сигналу, виражену в частках частоти дискретизації. Перший ФНЧ 11-го порядку має коефіцієнти 0,0003; -0,0578; -0,0559; 0,0850; 0,2974; 0,4000; 0,2974; 0,0850; -0,0559; -0,0578; 0,0003, та АЧХ цього фільтра містить значні коливання і далека від ідеальної.

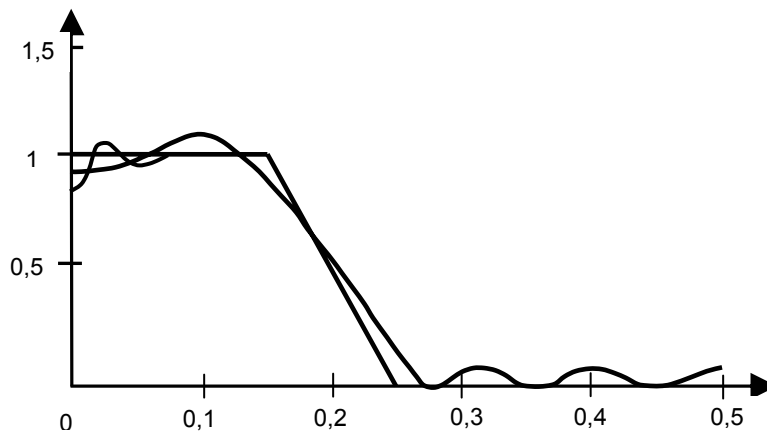


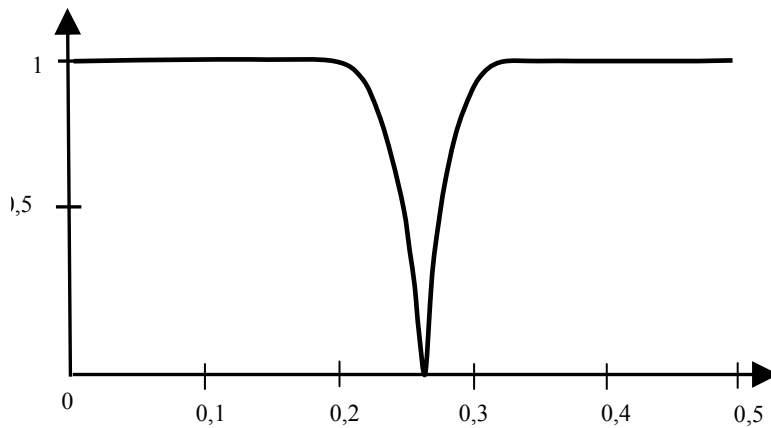
Рис. 3.2. Приклади АЧХ нерекурсивних фільтрів

Другий ФНЧ має порядок 101, АЧХ цього фільтра майже співпадає із заданою. Відхилення в зоні найнижчих частот зумовлені вибором неоптимального методу синтезу фільтра.

Якщо хоча б один з коефіцієнтів  $b_j$  не дорівнює нулю, то цифровий фільтр називають рекурсивним. У цьому разі на поточне

значення сигналу на виході впливають не лише значення вхідного сигналу, але і попередні значення сигналу на виході фільтра. Такі фільтри називають ще фільтрами з нескінченною імпульсною характеристикою (НІХ-фільтрами), оскільки вони «пам'ятають» усі попередні значення вхідного сигналу та сигналу на виході, і реакція рекурсивного фільтра на одиничний імпульс на вході теоретично триває нескінченно довго.

Амплітудно-частотну характеристику рекурсивного режекторного фільтра другого порядку, що має коефіцієнти  $a_0 = 0,9243$ ;  $a_1 = 0,1165$ ;  $a_2 = 0,9243$ ;  $b_1 = -0,1165$ ;  $b_2 = -0,8487$ , подано на рис. 3.3. Цей фільтр має частоту режекції, яка дорівнює 0,27 частоти дискретизації, і добротність, що становить 10. Режекторний фільтр має структуру, яку показано на рис. 3.1. Форма АЧХ залежить від вибору коефіцієнтів фільтра.



**Рис. 3.3.** АЧХ рекурсивного режекторного фільтра

Підвищуючи порядок цифрового фільтра, можна отримувати АЧХ із більш крутими спадами і таку, що відповідає поставленим

вимогам. Проте слід розуміти, що можливості поліпшення параметрів цифрових фільтрів не безмежні внаслідок похибок квантування та округлення, що виникають через скінченну кількість розрядів квантування сигналів, коефіцієнтів фільтра і скінченну розрядність пристроїв, що виконують обчислення за формулою (3.1). Вплив похибок квантування та округлення на результат фільтрації зростає з підвищенням порядку фільтра.

Використовуючи рекурсивні фільтри, можна отримати потрібні параметри АЧХ за більш низького порядку фільтра, ніж у разі застосування нерекурсивних фільтрів. Отже, обсяг ЗП і кількість арифметичних операцій можуть бути меншими. Але рекурсивні фільтри більш нестійкі, оскільки в них є зворотні зв'язки, і вони більше піддаються впливу похибок квантування та округлення.

Цифрову фільтрацію сигналу зазвичай здійснює процесор, що виконує множення і додавання, а також ЗП, у якому зберігаються відліки вхідного сигналу і сигналу на виході, що використовуються в обчисленнях за формулою (3.1). Цифровий фільтр може бути реалізований програмно на персональному комп'ютері, мікропроцесорі або мікроконтролері. Фільтрація звукових і телевізійних сигналів у реальному часі часто виконується за допомогою цифрових процесорів оброблення сигналів (ЦПОС) або спеціалізованих ВІС.

Фільтрацію, обчислену за формулою (3.1), називають фільтрацією в часовій області. Можливе також виконання фільтрації в частотній області за рахунок ДПФ вхідного сигналу. Отриманий набір коефіцієнтів ДПФ помножується на набір чисел, які являють собою частотну характеристику фільтра. Після цього виконується обернене ДПФ, тобто реалізується структурна схема оброблення, зображена на рис. 2.3. Завдяки використанню такого методу у



багатьох випадках досягають кращих результатів, ніж у разі фільтрації у часовій області. Але для виконання прямого і оберненого ДПФ телевізійних сигналів у реальному часі необхідні високопродуктивні обчислювальні засоби.

Слід пам'ятати, що кількість коефіцієнтів ДПФ, а отже, і кількість окремих частотних складових, які можна обробити, дорівнює кількості відліків у відрізку перетворюваного сигналу (див. підрозд. 2.2). Тому чим з більшою точністю належить реалізувати задану АЧХ, тим довші відрізки сигналу потрібно обробляти за допомогою ДПФ, а це, у свою чергу, зумовлює збільшення обсягу обчислень.

Фільтр, дія якого визначається співвідношенням (3.1), є лінійним. Якщо на його вхід подати суму декількох сигналів, то сигнал на виході буде сумою сигналів, отримуваних під дією кожного з вхідних сигналів окремо. Цю властивість лінійних фільтрів називають принципом суперпозиції.

Велике значення мають і нелінійні фільтри, для яких не виконується принцип суперпозиції. Одним з найвідоміших видів нелінійних фільтрів є медіанні фільтри. Принцип дії медіанного фільтра такий: для отримання кожного відліку сигналу на виході береться відповідний відлік вхідного сигналу і деяка кількість відліків, які йдуть перед ним та після нього. Значення цих відліків розташовують у порядку збільшення або зменшення. Далі значення центрального елемента отриманого набору чисел, який називають медіаною цього набору чисел, привласнюється відліку сигналу на виході.

Розглянемо приклад. Нехай маємо вхідний цифровий сигнал  
85, 86, 84, 84, 86, 48, 84, 86, 87, 85, 112, 144, 145, 143, 145, 146, 143, 144.

У сигналі є викид униз у 6-му відліку, а після 10-го відліку рівень

сигналу за два періоди дискретизації підвищується. Сигнал проходить через медіанний фільтр з шириною вікна у 5 відліків, яке переміщується зліва направо. Під час обчислення перших двох і останніх двох відліків сигналу на виході набір чисел доповнюється до 5 за рахунок повторення значень першого та останнього відліків вхідного сигналу відповідно. Хай центр вікна перебуває на 5-му відліку, у вікно потрапляють відліки із значеннями 86, 84, 84, 86, 48. Після впорядковування цей набір чисел набуває вигляду 48, 84, 84, 86, 86. Медіаною цього набору, а отже, значенням 5-го відліку сигналу на виході, буде число 84. Аналогічно обчислюються значення решти відліків, у результаті чого утворюється такий сигнал:

85, 85, 85, 84, 84, 84, 86, 85, 86, 87, 112, 143, 144, 145, 145, 144, 144, 144.

Короткий від'ємний викид у 6-му відліку усунено, помітно зменшилися дрібні флуктуації сигналу, додатний перепад і його крутизна не змінилися. Для порівняння наведемо результат оброблення того ж вхідного сигналу нерекурсивним ФНЧ 5-го порядку, у якому всі 5 коефіцієнтів дорівнюють 0,2. Такий ФНЧ формує відліки вихідного сигналу через усереднення за 5 відліками вхідного сигналу. Правило оброблення початкових і кінцевих відліків сигналу таке ж, як і для медіанного фільтра. З результату дії ФНЧ

85, 85, 85, 78, 77, 77, 78, 78, 91, 103, 115, 126, 138, 145, 144, 144, 144, 144.  
видно, що лінійний ФНЧ не усуває короткого викиду, а «розмазує» його на декілька відліків. Крім того, він зменшує крутий перепад у сигналі, що у разі телевізійного сигналу може призвести до погіршення чіткості зображення по горизонталі і розмивання меж об'єктів.

Таким чином, медіанний фільтр є ефективним засобом усунення

коротких імпульсних перешкод зі збереженням різких перепадів сигналу. Проте якщо короткий викид сигналу містить яку-небудь інформацію, то її буде втрачено, оскільки відрізнити перешкоду від корисної складової сигналу медіанний фільтр не може.

### 3.2. Просторова цифрова фільтрація

У сучасних телевізійних системах велику роль відіграють двовимірні або просторові цифрові фільтри. За просторової фільтрації відбувається двовимірна згортка вхідного зображення, поданого у вигляді матриці елементів, та матриці фільтра.

Нехай  $x(i, j)$  – початкове зображення;  $y(i, j)$  – зображення, що утворюється у результаті фільтрації;  $a(m, n)$  – матриця, яка визначає двовимірний фільтр і має розміри  $M$  по горизонталі і  $N$  по вертикалі, причому  $M$  і  $N$  мають бути непарними числами. Дію двовимірного нерекурсивного фільтра можна описати такою формулою

$$y(i, j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a(m, n) x\left[i + m - \frac{(M-1)}{2}, j + n - \frac{(N-1)}{2}\right], \quad (3.5)$$

де  $i$  – номер рядка;  $j$  – номер елемента в рядку.

Як видно з формули (3.5), кожний елемент зображення на виході фільтра являє собою суму елементів вхідного зображення, помножених на задані в матриці фільтра коефіцієнти. Таку операцію називають двовимірною згорткою. Кількість сусідніх урахуваних елементів вхідного зображення визначається розміром матриці фільтра.

Найпростіші просторові фільтри мають матрицю  $3 \times 3$ . Нижче приведено декілька прикладів фільтрів такого типу:

$$a) \begin{pmatrix} 0,1 & 0,1 & 0,1 \\ 0,1 & 0,2 & 0,1 \\ 0,1 & 0,1 & 0,1 \end{pmatrix}; \quad б) \begin{pmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{pmatrix}; \quad в) \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}.$$

Просторово-частотну характеристику (ПЧХ) для таких фільтрів у разі, коли матриця фільтру симетрична як відносно середнього стовпця і середнього рядка, так і відносно обох головних діагоналей, розраховують за формулою

$$H(\mu, \nu) = a_0 + 2a_1(\cos \mu + \cos \nu) + 2a_2[\cos(\mu + \nu) + \cos(\mu - \nu)], \quad (3.6)$$

де  $a_0$  – число в центральній комірці матриці фільтра;  $a_1$  – числа в середніх комірках верхнього і нижнього рядків і бічних стовпців;  $a_2$  – числа в кутових комірках;  $\mu = 2\pi f_x$ ;  $\nu = 2\pi f_y$ ;  $f_x$  і  $f_y$  – просторові частоти за координатами  $x$  і  $y$ .

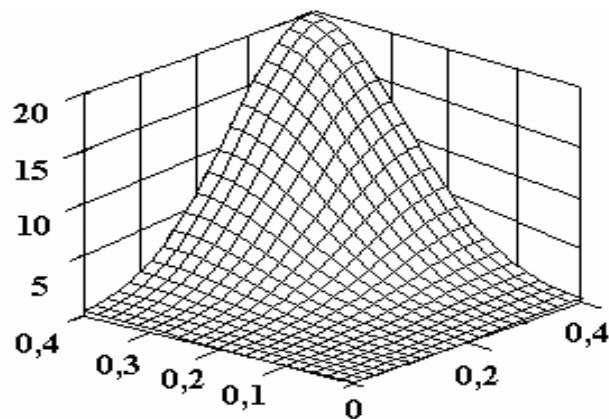
Фільтр (а) є фільтром нижніх просторових частот. Яскравість кожного елемента зображення на виході фільтра визначається через усереднення значень яскравості найближчих до нього сусідніх елементів початкового зображення. Сума всіх коефіцієнтів матриці дорівнює одиниці, що потрібно для збереження середнього значення яскравості зображення під час фільтрації. Просторові частоти  $f_x$  і  $f_y$  вимірюють у частках просторової частоти дискретизації (тобто величини, оберненої відстані між сусідніми відліками) за відповідною координатою, і кожна з них змінюється в діапазоні  $0 \dots 0,5$ .

Такий фільтр зменшує помітність шуму на зображенні, але при цьому зменшуються різкі перепади яскравості, погіршується передавання меж об'єктів, можуть зникати дрібні деталі зображення.

Фільтр (б) є фільтром верхніх просторових частот. Значення

сусідніх елементів віднімаються від центрального елемента зображення. Сума всіх коефіцієнтів дорівнює одиниці. У результаті фільтрації підкреслюються межі об'єктів, підвищується різкість зображення, але одночасно стає більш помітним шум. Просторово-частотну характеристику цього фільтра подано на рис. 3.4.

Фільтр ( $\nabla$ ) належить до операторів Лапласа. Сума коефіцієнтів, на відміну від попередніх фільтрів, дорівнює нулю, внаслідок чого середній рівень яскравості зображення під час фільтрації не зберігається. Призначення фільтрів, що реалізують оператор Лапласа, виділяти лише межі або контури об'єктів, тобто лінії, за якими відбувається різка зміна яскравості. Ділянки з постійною яскравістю такий фільтр перетворює на темні ділянки зображення.



**Рис. 3.4.** ПЧХ фільтра верхніх просторових частот

Двовимірну фільтрацію широко використовують під час оброблення зображень на ЕОМ. При цьому окремий кадр телевізійного сигналу або зображення, отримане яким-небудь іншим способом, вводиться в ЕОМ і обробляється з метою зменшення шумів, виділення меж об'єктів, підвищення чіткості тощо.

Двовимірна фільтрація в реальному часі зі швидкістю телевізійної розгортки зазвичай реалізується за допомогою паралельних обчислювальних пристроїв, спеціальних процесорів згортки та інших апаратних засобів.

Крім фільтрації за допомогою згортки цифрового сигналу з матрицею, що визначає фільтр, можливе виконання фільтрації в частотній області через двовимірне ДПФ вхідного зображення, множення коефіцієнтів ДПФ на числа, що являють собою ПЧХ фільтра, та обернене двовимірне ДПФ.

Для фільтрації зображень застосовуються двовимірні медіанні фільтри з урахуванням сусідніх елементів не лише по горизонталі, але і по вертикалі. Медіанний фільтр добре усуває точкові перешкоди і шуми, зберігаючи при цьому різкі перепади яскравості на межах об'єктів. Проте такий фільтр може придушити у зображенні дрібні деталі, розміри яких менші від розмірів вікна фільтра.

### **3.3. Застосування цифрової фільтрації для зменшення шумів**

Основними джерелами шуму в телевізійному сигналі є перетворювач світло-сигнал і вхідний каскад телевізійного приймача. Перешкоди, що виникають у каналі зв'язку, також можуть впливати на зображення. Ще одне джерело шумів у телевізійному сигналі – апаратура магнітного відеозапису, з якої цей сигнал відтворюється.

Шуми, що впливають на зображення, перебувають в одній смузі частот із сигналами яскравості й кольору. Щоб зменшити за допомогою частотних фільтрів рівень шумів, не погіршуючи при цьому інших параметрів якості зображення, потрібно або

використати розбіжності у структурі спектрів корисного сигналу й шуму, або ж допускати зменшення разом із шумом деяких частотних складових телевізійного сигналу так, щоб при цьому погіршення якості зображення, яке сприймається глядачем, було в допустимих межах.

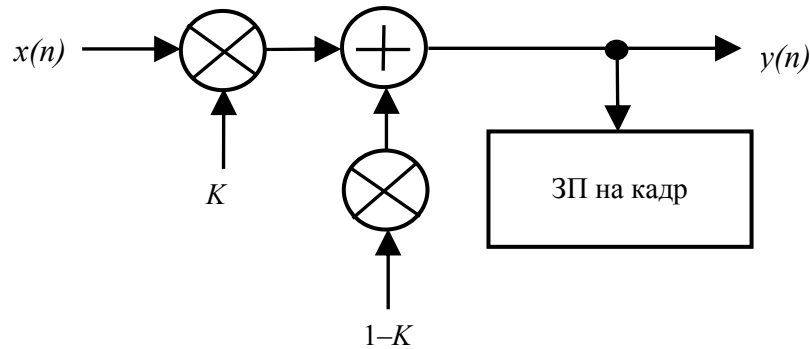
Розглянемо можливі варіанти виконання фільтрації для зменшення помітності шуму.

Найпростіший варіант – обмеження смуги частот телевізійного сигналу, оскільки відомо, що чим вужче смуга частот, тим менша потужність шуму. Але цей варіант не може дати потрібного результату, тому що разом із шумами будуть зменшені високочастотні складові сигналу, що призведе до погіршення чіткості по горизонталі та різкості меж об'єктів.

Небажано також застосовувати простий просторовий ФНЧ, який зменшує високі просторові частоти. Такий фільтр погіршить чіткість зображення як по горизонталі, так і по вертикалі. Медіанний фільтр добре усуває точкові перешкоди, але оброблення зображення з його допомогою може спричинити втрати на зображенні контрастних деталей малих розмірів.

Значно кращі результати дає цифрова фільтрація у часі окремо для кожного елемента зображення в кадрі, суть якої полягає в тому, що вхідний цифровий сигнал  $x(n)$  (рис. 3.5) помножується на коефіцієнт  $K$  ( $0 \leq K \leq 1$ ) і додається до сигналу на виході, затриманого на 1 кадр і помноженого на  $(1 - K)$ . Затримка на кадр відповідає затримці на  $N_{\text{кадр}} = f_d \cdot T_{\text{кадр}}$  елементів зображення, де  $f_d$  – частота дискретизації,  $T_{\text{кадр}}$  – період кадрової розгортки. Якщо відповідно до Рекомендації 601  $f_d = 13,5$  МГц, а  $T_{\text{кадр}} = 40$  мс, то  $N_{\text{кадр}} = 540000$ . При цьому сигнал кожного елемента зображення чергового кадру на вході фільтра буде додаватися до сигналу

відповідного елемента попереднього кадру на виході фільтра, тобто забезпечується усереднення за часом сигналів кожного елемента зображення.



**Рис. 3.5.** Рекурсивний фільтр із затримкою на кадр

Зменшення помітності шуму відбувається через те, що в нерухомих зображеннях або у зображеннях, які змінюються повільно, корисні сигнали елементів зображення повторюються від кадру до кадру, у той час як значення напруги шуму кожного елемента зображення від кадру до кадру змінюються випадково. Тому в результаті усереднення шум буде послаблятися, а корисний сигнал залишиться.

Дія фільтра описується співвідношенням

$$y(n) = K x(n) + (1 - K) y(n - N_{\text{кадр}}). \quad (3.7)$$

Таку цифрову фільтрацію у часі виконує рекурсивний фільтр порядку  $N_{\text{кадр}}$ . Залежності від параметра  $K$  значень коефіцієнта зменшення шуму  $K_{\text{ш}}$  і кількості кадрів  $N_e$ , що характеризує інерційність фільтра, наведено у табл. 3.1. За  $N_e$  кадрів після вимкнення сигналу на вході фільтра сигнал на виході фільтра зменшується приблизно у  $e = 2,72$  рази.



**Таблиця 3.1.** Параметри рекурсивного фільтра

$K$	0,5	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001
$K_{ш}$	1,73	3,00	4,36	6,25	9,95	14,1	20,0	31,6	44,7
$K_{ш}, \text{дБ}$	4,8	9,5	12,8	15,9	20,0	23,0	26,0	30,0	33,0
$N_e$	2	5	10	20	50	100	200	500	1000

Амплитудно-частотна характеристика такого фільтра має вигляд «гребінки» зі смугами пропускання на частотах, кратними частоті кадрів 25 Гц. Такі фільтри називають *гребінчастими*. Ширина піків АЧХ залежить від значення коефіцієнта  $K$ : чим менше  $K$ , тим вужчі утворюються смуги пропускання. У разі передавання нерухомого зображення у спектрі телевізійного сигналу будуть лише гармоніки частоти 25 Гц, тому що кожний елемент зображення періодично повторюється з частотою кадрів. У той же час, шум має неперервний спектр. Гребінчастий фільтр пропускати гармоніки сигналу та істотно послаблюватиме шум. Чим вужчі піки АЧХ, тим менша частка потужності шуму пройде на вихід фільтра.

У разі передавання зображень, що змінюються, дія такого фільтра створюватиме накладення зображень попередніх кадрів на наступні, що зумовить розмазування зображень рухомих об'єктів та інші спотворення. Щоб запобігти цим негативним явищам, такі фільтри роблять адаптивними до величини змін зображення від кадру до кадру.

Структурну схему адаптивного фільтра для зниження шумів, що входить до складу BIC SAA4990H виробництва фірми «PHILIPS», наведено на рис. 3.6. Для роботи фільтра необхідна додаткова BIC ЗП на одне поле телевізійного сигналу. Затримка здійснюється на

період полів, тому у співвідношенні (3.7) замість кількості елементів у кадрі  $N_{\text{кадр}}$  буде кількість елементів у полі  $N_{\text{пол}}$ , а піки АЧХ фільтра будуть знаходитися на частотах, кратних 50 Гц. На відміну від рекурсивного фільтра (рис. 3.5), у адаптивному фільтрі алгоритм (3.7) реалізується за допомогою двох суматорів Сум1 і Сум2 та одного помножувача П1. Параметр  $K$  виробляється детектором руху ДР, що містить фільтри  $\Phi 1$  і  $\Phi 2$ , обмежувач Об, помножувач П2 на постійний коефіцієнт і нелінійний перетворювач (НП). Всі ці вузли цифрові.

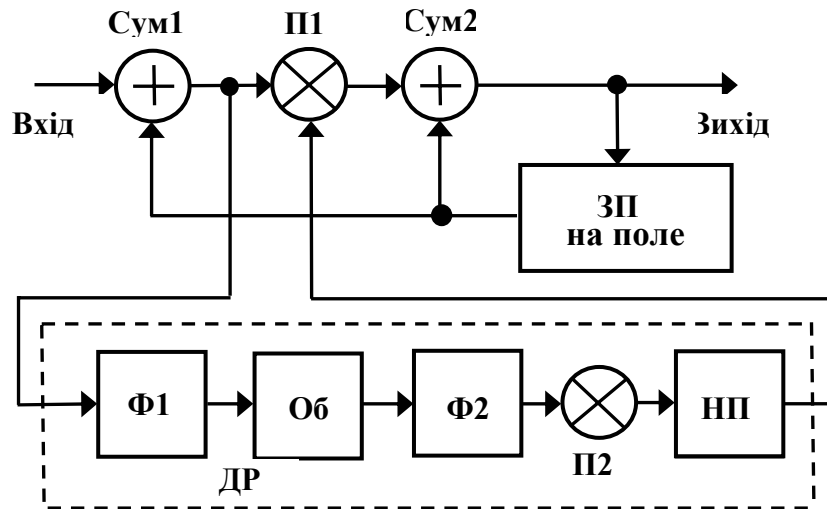


Рис. 3.6. Адаптивний фільтр

На вхід детектора руху (ДР) із суматора Сум1 надходить різниця вхідного сигналу  $x(n)$  і сигналу на виході  $y(n - N_{\text{пол}})$ , затриманого на поле. Фільтри  $\Phi 1$  і  $\Phi 2$  – нерекурсивні режекторні фільтри відповідно 2-го й 4-го порядків. Вони зменшують частотні складові різницевого сигналу в діапазоні 0,25...0,5 частоти дискретизації. В обмежувачі Об рівень цього сигналу обмежується зверху, а в

помножувачі П2 – помножується на постійний коефіцієнт із набору 1, 2, 4, 8. Нелінійний перетворювач НП має передавальну характеристику, що збільшується від нуля спочатку швидко, потім повільніше і, нарешті, зупиняється на рівні  $K = 1$ .

Детектор руху ДР обробляє зміни яскравості елементів зображення відносно відповідних елементів зображення попереднього поля. Якщо зміна яскравості мала, то для такого елемента зображення параметр  $K$  встановлюється малим, і цей елемент зображення у вихідному сигналі отримується через усереднення відповідних елементів зображення з багатьох попередніх полів. У результаті цього шум послаблюється значною мірою (табл. 3.1).

Якщо ж зміни яскравості елементів зображення більші, то параметр  $K$  стає ближчим до одиниці, у результаті чого кількість полів, по яким виконується усереднення, зменшується. Коли  $K$  досягає одиниці, усереднення та ослаблення шуму не відбуваються. Змінюючи постійний коефіцієнт у помножувачі П2, можна збільшувати або зменшувати ступінь ослаблення шумів, однак чим більше ослаблення шуму, тим більша інерційність фільтра.

BIC SAA4990H містить два фільтри для зменшення шумів: перший для сигналу яскравості, а другий – для кольорорізницевих сигналів, які обробляються по черзі через елемент відповідно до формату 4:1:1. Детектор руху наявний лише тільки у фільтрі для сигналу яскравості, а параметр  $K$ , який виробляється в цьому фільтрі, надходить також і на фільтр кольорорізницевих сигналів.

Адаптація до змін яскравості для кожного елемента зображення майже не залежить від змін яскравості інших елементів. Невелика кореляція коефіцієнтів  $K$  для сусідніх у рядку елементів створюється фільтрами Ф1 і Ф2.

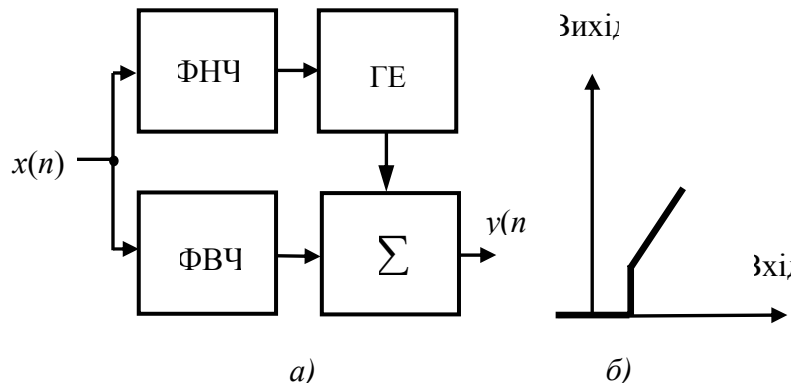
Слід зазначити, що хороші результати досягаються за допомогою адаптивних фільтрів завдяки тому, що принцип їх роботи узгоджений з властивостями зору. Людина менше помічає шум на рухомих зображеннях і на ділянках з різкими перепадами яскравості, де фільтр не діє, і більше сприймає шум саме на нерухомих та гладких ділянках зображень, на яких і працює фільтр.

Інший метод зменшення помітності шуму на зображенні, який використовують на практиці, ґрунтується на одночасному застосуванні лінійної фільтрації й нелінійної граничної обробки високочастотних складових телевізійного сигналу. Вхідний цифровий сигнал  $x(n)$  (рис. 3.7, а) за допомогою ФНЧ і ФВЧ розділяється на низькочастотну й високочастотну складові. Високочастотна складова проходить через граничний елемент ГЕ, передавальну характеристику якого подано на рис. 3.7, б, після чого додається до низькочастотної складової.

Якщо високочастотна складова телевізійного сигналу має досить високий рівень, наприклад на ділянках зображення, які містять контури об'єктів, контрастні деталі малих розмірів тощо, то вона не змінюється, тому такі ділянки зображення не спотворюються. Якщо ж високочастотна складова має низький рівень, тобто зумовлена шумом на фоні гладких ділянок зображення, то вона не пропускається граничним елементом і, як результат, шум на гладких ділянках зображення зменшується.

Фільтр із граничною обробкою не містить ЗП на поле й навіть на рядок, його дія практично безінерційна, однак у нього є істотні недоліки. Текстура поверхонь об'єктів буде зменшуватися, як і шуми. У той же час низькочастотні випадкові перешкоди проходять через фільтр. Такий фільтр реалізований, наприклад, у BIC SAA7165 виробництва фірми «PHILIPS», рівень порога у якій

задається за допомогою шини  $I^2C$  і може набувати 4-х значень, а граничній обробці піддається лише сигнал яскравості.



**Рис. 3.7.** Фільтр із граничною обробкою (а)  
і передавальна характеристика граничного елемента (б)

### 3.4. Перетворення розгортки

Телевізійне зображення є функцією двох просторових координат і часу. У процесі розгортки здійснюється дискретизація телевізійного зображення у часі з частотою кадрів і за вертикальною координатою з визначеною кількістю рядків. Під час перетворення телевізійних сигналів у цифрову форму здійснюється також дискретизація по горизонталі. Перетворення розгортки у телевізійних системах у більшості випадків полягає у змінюванні параметрів дискретизації за часом (частоти кадрів і полів) або за вертикальною координатою (кількості рядків). Використання цифрової техніки зробило можливим виконання різних видів перетворення розгортки [5], що дає змогу вийти на новий рівень якості телевізійного зображення, залишаючись у рамках чинних

стандартів розкладення зображення.

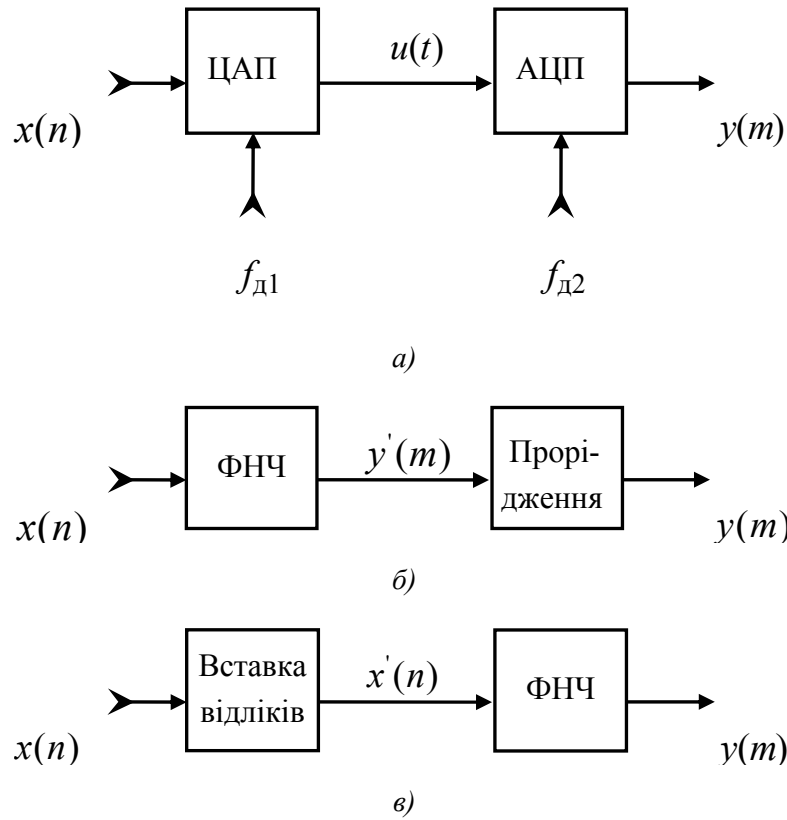
Прикладами перетворення розгортки можуть бути перетворення сигналу NTSC (525 рядків, 30 кадрів за секунду) на сигнал PAL або SECAM (625 рядків, 25 кадрів за секунду) або навпаки, підвищення частоти кадрової розгортки в телевізорі до 100 Гц, формування зменшеного зображення пристроєм «кадр у кадрі», перетворення телевізійного сигналу з черезрядковою розгорткою в сигнал із прогресивною розгорткою тощо.

Розглянемо загальний принцип змінювання частоти дискретизації за однією координатою, яке будемо називати *передискретизацією* (англ. resampling). Щоб виконати передискретизацію, необхідно цифровий сигнал  $x(n)$ , що має першу частоту дискретизації  $f_{d1}$ , перетворити в аналоговий сигнал  $u(t)$ , а потім цей аналоговий сигнал знову перетворити у цифровий сигнал  $y(n)$ , але вже з іншою частотою дискретизації  $f_{d2}$  (рис. 3.8, а).

Відзначимо, що у разі змінювання кількості рядків під частотою дискретизації розуміють просторову частоту по вертикалі, обернено пропорційну кількості рядків.

На практиці всі операції виконують над сигналами у цифровій формі, не повертаючись до аналогового подання, щоб уникнути додаткових спотворень сигналів і спростити апаратуру. Під час цифро-аналогового перетворення слід використовувати ідеальний ФНЧ, що усуває всі частоти, які вищі за половину частоти дискретизації. У процесі передискретизації у цифровій формі функції цього ФНЧ виконує цифровий фільтр. Передискретизація полягає в обчисленні значень відліків сигналу на виході  $y(n)$  за відліками вхідного сигналу  $x(n)$ . ФНЧ при цьому не лише імітує виконання цифро-аналогового перетворення вхідного сигналу, але й

має забезпечувати виконання умови Котельникова для дискретизації на виході. У разі зниження частоти дискретизації (downsampling) частота зрізу ФНЧ повинна дорівнювати  $f_{д2}/2$ , а у разі підвищення частоти дискретизації (upsampling) –  $f_{д1}/2$ .



**Рис. 3.8.** Загальна структура зміни частоти дискретизації (а) і структурні схеми зниження (б) і підвищення (в) частоти дискретизації

Порівняно просто реалізується зміна частоти дискретизації в цілу кількість разів. Для зниження частоти дискретизації у  $k$  разів

сигнал  $x(n)$  пропускається через ФНЧ із частотою зрізу, рівною  $f_{д1}/2$ . Після цього виконується проріджування (децимація) у  $k$  разів, у результаті чого у сигналі на виході  $y(m)$  залишається кожний  $k$ -й відлік профільтованого вхідного сигналу  $x'(n)$  (рис. 3.8, б). Приклад такого перетворення подано на рис 3.9, де кожен відлік цифрового сигналу зображено у вигляді стовпця.

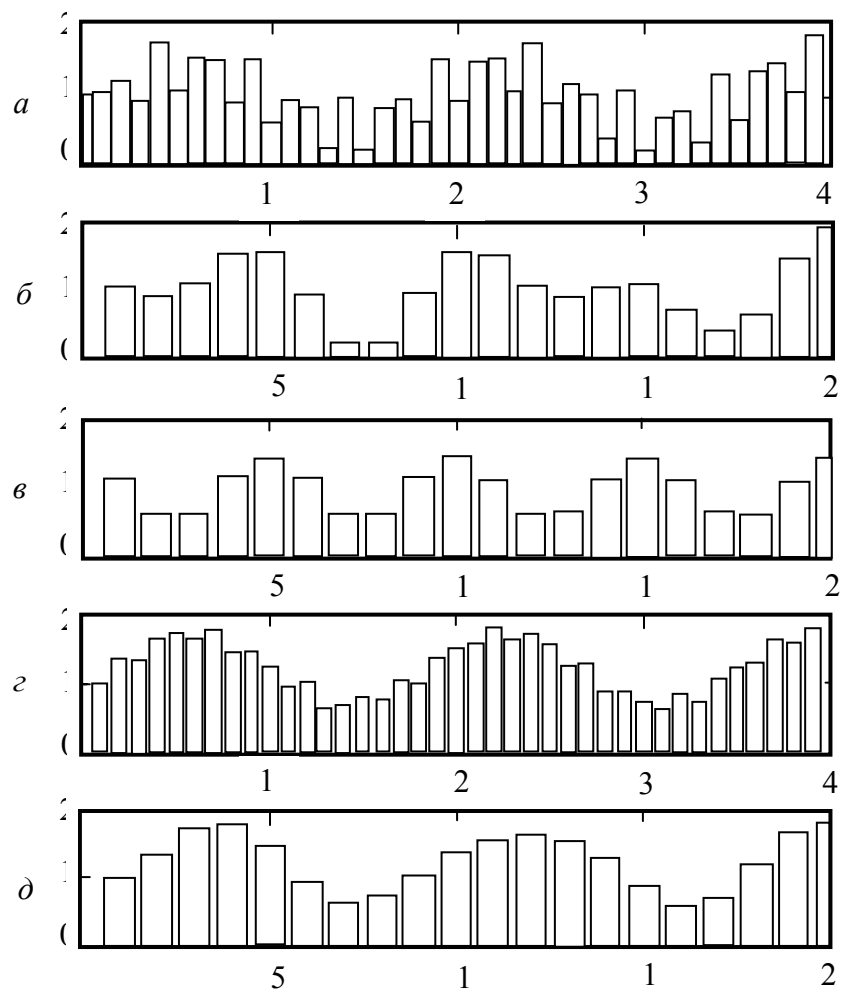
Вхідний сигнал (рис. 3.9, а) має вигляд

$$x(n) = 1 + 0,45 \cos(0,12 \pi n) + 0,45 \cos(0,8 \pi n).$$

Другий і третій доданки мають частоти, які дорівнюють відповідно 0,06 і 0,4 частоти дискретизації. У разі зменшення частоти дискретизації у два рази відкиданням кожного другого відліку частоти цих доданків стають відповідно 0,12 та 0,8 від нової частоти дискретизації  $f_{д2}$ . Частота 0,8 неприпустима, тому що у дискретизованому сигналі не має бути частот, більших від половини частоти дискретизації. У результаті порушення цієї умови в сигналі, отриманому завдяки проріджуванню вхідного сигналу у 2 рази (рис. 3.9, б), виникає хибна частотна складова. На рис. 3.9, б наявності хибної складової не видно, тому, щоб показати, що вона є, на рис. 3.9, в подано результат проріджування у два рази сигналу  $x(n) = 1 + 0,45 \cos(0,8 \pi n)$ , у якому залишилася лише неприпустима частота. Після проріджування з'явилася частота, яка дорівнює 0,2 від частоти дискретизації на виході  $f_{д2}$ , чого насправді не має бути. Цю частотну складову можна помітити і на рис. 3.9, б.

Щоб не виникало хибних частотних складових, у вхідному сигналі перед проріджуванням слід усунути всі частоти, що перевищують чверть вхідної частоти дискретизації.





**Рис. 3.9.** Зниження частоти дискретизації:

$a$  – вхідний сигнал;  $б$  – проріджування вхідного сигналу у 2 рази;  
 $в$  – проріджування сигналу  $x(n) = 1 + 0,45 \cos(0,8 \pi n)$  у 2 рази;  $г$  –  
 результат оброблення вхідного сигналу нерекурсивним цифровим  
 фільтром 11-го порядку;  $д$  – проріджування у 2 рази обробленого  
 вхідного сигналу

Нерекурсивний фільтр ФНЧ з ідеальною АЧХ і частотою зрізу, що дорівнює  $f_{д1}/4$ , визначається співвідношенням

$$x'(n) = \sum_{k=0}^{\infty} x(n-k) \frac{\sin(\pi k / 2)}{\pi k / 2}. \quad (3.8)$$

Такий фільтр має враховувати нескінченну кількість попередніх відліків вхідного сигналу, що реалізувати не можливо. Під час обчислень залишають кілька перших доданків у формулі (3.8), отримуючи ФНЧ, що апроксимує з деякою похибкою ідеальний ФНЧ. Результат оброблення вхідного сигналу нерекурсивним цифровим фільтром 11-го порядку, що містить 11 перших доданків за формулою (3.8), де складову з частотою  $0,4f_{д1}$  усунено, а складова з частотою  $0,06 f_{д1}$  залишилася, наведено на рис. 3.9, з. Після проріджування отриманого сигналу утворюється складова з частотою  $0,12 f_{д2}$ , період якої трохи більший за 8 періодів відліків, як і має бути (рис. 3.9, д). Цей приклад ще раз доводить необхідність дотримуватися умови Котельникова.

Для підвищення частоти дискретизації у  $k$  разів у сигналі  $x(n)$  між кожними двома сусідніми відліками вставляється по  $k-1$  нульових відліків, а потім отриманий сигнал пропускається через ФНЧ із частотою зрізу, рівною  $f_{д1}/2$  (рис. 3.8, в). Вибір частоти зрізу зумовлений тим, що у сигналі на виході не можуть з'явитися складові з більш високими частотами, ніж у вхідному сигналі. Приклад підвищення частоти дискретизації показано на рис. 3.10.

Вхідний сигнал (рис. 3.10, а) має вигляд  $x(n) = 1 + 0,45 \cos(0,24 \pi n)$ , а частота його змінної складової дорівнює  $0,12 f_{д1}$ . Для підвищення частоти дискретизації у 4 рази у проміжки між відліками вхідного сигналу вставляється по три нульових відліки (рис. 3.10, б). Потім здійснюється інтерполяція – присвоєння новим відлікам значень за певним законом.

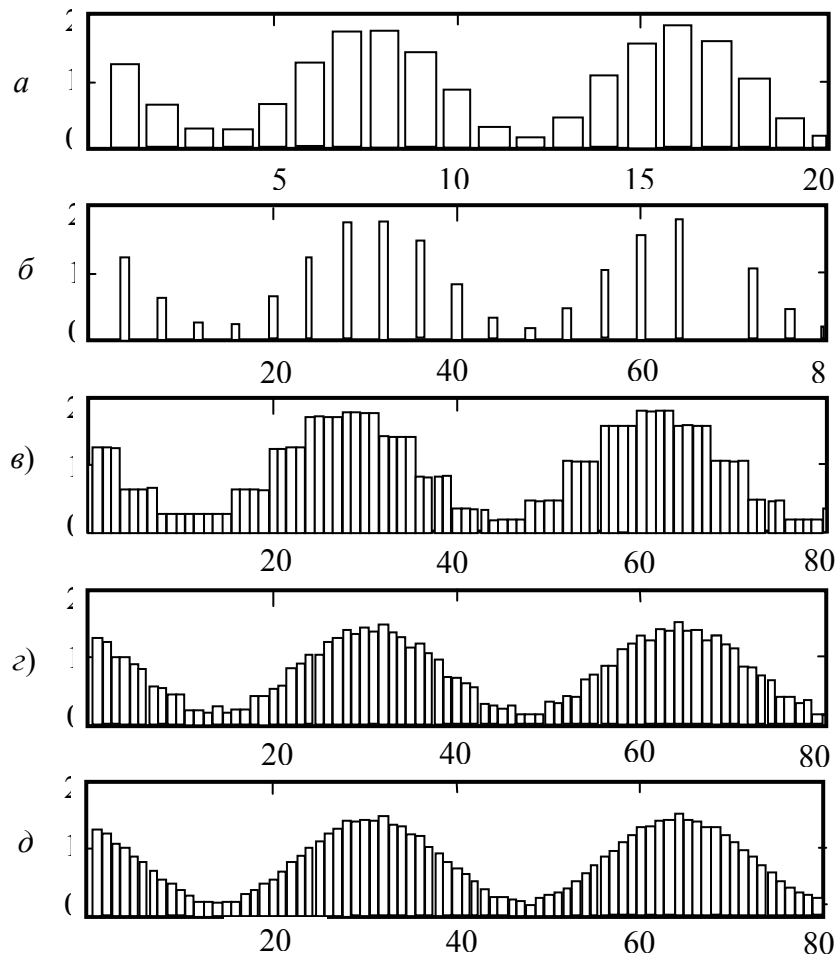
Найпростіший і найбільш поширений варіант інтерполяції – східчаста інтерполяція (рис. 3.10, *в*), але вона не забезпечує потрібної форми сигналу на виході. Значно кращий результат дає інтерполяція із застосуванням нерекурсивного цифрового фільтра, який апроксимує ідеальний ФНЧ із частотою зрізу  $f_{д1}/2$  та описується співвідношенням

$$y(m) = \sum_{k=-M}^M y'(m-k) \frac{\sin(\pi k/4)}{\pi k/4}. \quad (3.9)$$

Цифровий сигнал піддається дії ФНЧ, котрий має частоту дискретизації  $f_{д2}$  та частоту зрізу АЧХ 0,125 відносно  $f_{д2}$ . Для обчислення проміжних відліків сигналу на виході слід урахувати як попередні, так і наступні відліки вхідного сигналу, тому фільтр створює затримку сигналу на  $M$  відліків. Результати інтерполяції з застосуванням нерекурсивних фільтрів відповідно 21-го ( $M = 10$ ) і 41-го ( $M = 20$ ) порядків зображено на рис. 3.10, *з*, *д*.

Поєднуючи два описаних методи, можна змінювати частоту дискретизації у  $k_1/k_2$  разів, де  $k_1$  і  $k_2$  – цілі числа. Для цього треба спочатку збільшити частоту дискретизації у  $k_1$  разів, а потім зменшити її у  $k_2$  разів. На практиці обидві операції виконують в одному цифровому фільтрі, який відразу обчислює значення відліків вихідного сигналу з частотою дискретизації  $f_{д2} = f_{д1} k_1/k_2$ . Наприклад, під час перетворення телевізійного сигналу з кількістю рядків 525 у телевізійний сигнал з кількістю рядків 625 змінюється просторова частота дискретизації по вертикалі у 25/21 разів.

Якщо потрібно змінити не лише кількість рядків, але й кількість елементів зображення в рядку у процесі переходу від зображення високої до зображення звичайної чіткості або навпаки, то варто використовувати просторовий (двовимірний) ФНЧ.



**Рис. 3.10.** Підвищення частоти дискретизації:

*a* – вхідний сигнал; *б* – вставка в проміжки між відліками вхідного сигналу додаткових відліків; *в* – східчаста інтерполяція значень додаткових відліків; інтерполяція додаткових відліків із застосуванням нерекурсивних фільтрів 21-го (*г*) та 41-го (*д*) порядків

Апроксимацію ідеальних ФНЧ за допомогою нерекурсивних фільтрів досить високого порядку застосовують у студійній телевізійній апаратурі. Для перетворення розгортки у більш простих телевізійних системах використовують спрощені алгоритми фільтрації, щоб зменшити обсяг обчислень та знизити складність і вартість апаратури.

Один з найбільш відомих прикладів перетворення розгортки в телевізійних приймачах – отримання зображення зменшеного формату у пристроях «кадр у кадрі» (Picture in Picture – PIP). Кількість елементів зображення в рядку і кількість рядків зазвичай зменшуються у 3 або у 4 рази порівняно з нормальним телевізійним зображенням.

У PIP-процесорах SDA9188-3X й SDA9288X виробництва фірми «SIEMENS» кожен елемент зменшеного зображення отримують за допомогою усереднення  $3 \times 3$  або  $4 \times 4$  елементів вхідного зображення. Частотна характеристика таких фільтрів не ідеальна та не забезпечує повного усунення частот, що перевищують половину частоти дискретизації за кожною координатою, однак на реальних зображеннях помітних спотворень не створює.

Велике значення для підвищення якості зображення мають збільшення частоти вертикальної розгортки і перетворення черезрядкової розгортки у прогресивну. Це пов'язано з тим, що у разі збільшення розмірів і яскравості екранів кінескопів стають більш помітними мерехтіння із частотою 50...60 Гц і спотворення, які створюються черезрядковою розгорткою.

Для вирішення обох завдань доцільно змінити частоту дискретизації телевізійного зображення у часі. Розглянемо деякі варіанти виконання таких перетворень.

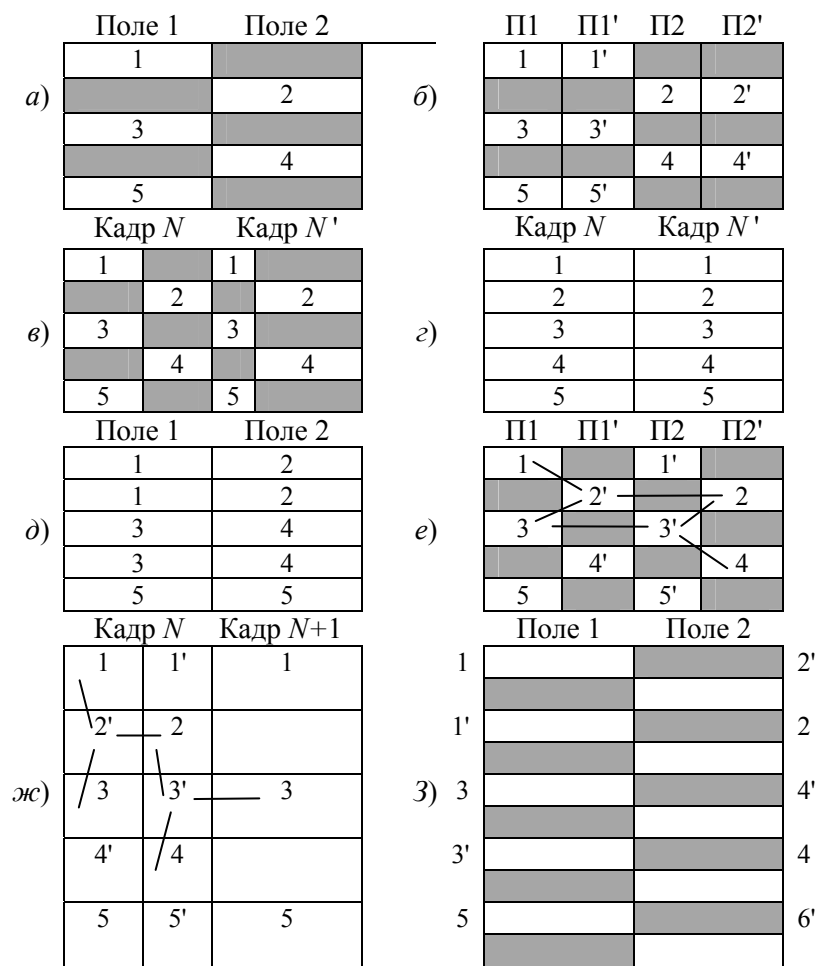
Положення рядків у першому і другому полях за звичайної

черезрядкової розгортки показано на рис. 3.11, а; рядки пронумеровано в порядку їх розташування по вертикалі; проміжки між рядками заштриховано.

Найпростіший варіант подвоєння частоти полів полягає в тому, що кожне поле записується в ЗП і зчитується звідти два рази з подвоєною швидкістю (рис. 3.11, б). З поля 1 виходять два однакових поля П1 і П1', а з поля 2 – П2 і П2', що дає можливість майже усунути помітність мерехтінь на об'єктах великих розмірів, але мерехтіння горизонтальних і похилих ліній із товщиною, порівняну з шириною рядків, залишаються, тому що зміна парних і непарних рядків на екрані відбувається з попередньою частотою.

Інший варіант полягає в подвоєнні цілих кадрів (рис. 3.11, в). Кадр запам'ятовується, а потім відтворюється два рази з подвоєною частотою розгортки, у результаті чого утворюються два кадри: кадр  $N$  і кадр  $N'$ , що складаються з відповідних полів. У цьому разі ефект мерехтінь тонких ліній стане менш помітним. Однак цей варіант має недолік: у вхідному телевізійному зображенні парні рядки в часі відтворюються після непарних рядків відповідного кадру, а після перетворення розгортки виявляється, що непарні рядки кадру  $N'$  відтворюються після парних рядків кадру  $N$ , хоча насправді вони відповідають більш ранньому інтервалу часу. Це може призвести до спотворень зображень рухомих об'єктів.

Один із простих варіантів перетворення черезрядкової розгортки в прогресивну полягає в запам'ятовуванні обох полів кадру та їх спільному відтворенні два рази (рис. 3.11, г). Цей метод добре працює на нерухомих зображеннях. Якщо ж у зображенні є рухомі об'єкти, то виникнуть спотворення, тому що під час об'єднання рядків двох полів відбувається сполучення в часі положень об'єктів, які насправді не були одночасними.



**Рис. 3.11.** Варіанти подвоєння частоти полів і перетворення черезрядкової розгортки у прогресивну:

*a* – черезрядкова розгортка; *б* – подвоєння частоти полів; *в* – подвоєння частоти кадрів; *г*, *д* – перетворення черезрядкової розгортки у квазіпрогресивну; подвоєння кадрової частоти (*е*); перетворення черезрядкової розгортки у прогресивну (*жс*) та подвоєння кількості рядків за черезрядкової розгортки (*з*)

Відзначимо, що розгортку, яка виникає в результаті такого перетворення, називають квазіпрогресивною, тому що у разі повноцінної прогресивної розгортки всі рядки кадру повинні передаватися підряд.

Інший варіант перетворення черезрядкової розгортки у прогресивну полягає у подвоєнні кожного рядка в першому та другому полях (рис. 3.11, д). Цей варіант також має недоліки, тому що, по-перше, зберігається мерехтіння тонких горизонтальних ліній, а по-друге, погіршується чіткість по вертикалі.

Зазначені приклади свідчать про те, що прості методи перетворення розгортки у часі мають недоліки, для подолання яких слід застосовувати більш досконалі алгоритми комбінування рядків парних і непарних полів. Кілька поліпшених варіантів перетворення розгортки реалізуються із застосуванням BIC SAA4990H фірми «PHILIPS».

У разі подвоєння частоти вертикальної розгортки кадр перетворюється на два кадри, перший з яких містить поля П1 і П2', а другий – поля П1' і П2 (рис. 3.11, е). Але, на відміну від варіанта подвоєння кадрів згідно з рис. 3.11, в, рядки в додаткових полях П2' і П1' формуються за допомогою інтерполяції рядків першого і другого полів вхідного кадру. Інтерпольовані рядки позначено штрихами (1', 2' і т. д.). Відрізки прямих на рис. 3.11, е показують, які саме рядки вихідного кадру використовуються для інтерполяції рядків додаткових полів.

Зупинимось на особливостях виконання інтерполяції у BIC SAA4990H. Для отримання додаткових відліків у проміжках між відліками вихідного сигналу, тобто для інтерполяції, потрібно використовувати цифровий ФНЧ, який має досить високий порядок та близький до ідеального. Збільшення порядку ФНЧ на одиницю



вимагає додавання ще одного ЗП на поле, а також збільшення кількості арифметичних операцій. Тому в SAA4990H замість лінійного ФНЧ, що описується співвідношенням (3.8) або (3.9), застосовується медіанний фільтр. Кожен відлік у рядку 2' отримується як медіана значень відповідних відліків з рядків 1, 2 і 3, тобто як проміжний за рівнем сигналу із цих трьох відліків. Аналогічно формуються відліки в інших інтерпольованих рядках. Завдяки цьому зменшуються спотворення рухомих об'єктів.

В іншому режимі роботи ВІС SAA4990H забезпечує перетворення черезрядкової розгортки у прогресивну (рис. 3.11, ж), кількість рядків у кожному полі під час перетворення подвоюється. Додаткові рядки 1', 2' і т. д. формуються завдяки медіанній фільтрації, як і у попередньому варіанті. Відрізки прямих на рис. 3.11, ж показують, які рядки вхідного телевізійного зображення беруть участь у формуванні додаткових рядків першого і другого полів, тобто для отримання додаткових рядків другого поля використовуються непарні рядки з першого поля наступного кадру.

Ще один режим роботи SAA4990H – подвоєння кількості рядків за умови збереження черезрядкової розгортки (рис. 3.11, з). У цьому режимі 625 рядків перетворюються на 1250, тобто відбувається імітація зображення високої чіткості. Додаткові рядки утворюються за допомогою інтерполяції з використанням медіанного фільтра.

Слід зазначити, що завдання перетворення розгортки зберігають актуальність і для приймачів нових систем цифрового телебачення, тому що згідно чинних стандартів розкладення зображення у них зберігається передавання з черезрядковою розгорткою і частотою полів 50 Гц.

## **Контрольні запитання**

- 1. Напишіть формулу, що описує дію цифрового фільтра.*
- 2. Наведіть структурну схему цифрового фільтра.*
- 3. Що таке передавальна та імпульсна характеристики цифрового фільтра?*
- 4. Дайте характеристику нерекурсивного цифрового фільтра.*
- 5. Дайте характеристику рекурсивного цифрового фільтра.*
- 6. Дайте характеристику медіанного цифрового фільтра.*
- 7. Поясніть принцип роботи просторових двовимірних фільтрів.*
- 8. Поясніть принцип роботи рекурсивного фільтра для зменшення шумів на зображенні з затримкою на кадр.*
- 9. Поясніть принцип роботи адаптивного фільтра для зменшення шумів на зображенні.*
- 10. Як відбувається зменшення частоти дискретизації під час перетворення розгортки?*
- 11. Як відбувається збільшення частоти дискретизації під час перетворення розгортки?*
- 12. Як відбувається збільшення частоти зміни кадрів і полів?*
- 13. Які методи зменшення спотворень рухомих об'єктів застосовуються під час перетворення розгортки?*
- 14. Як перетворюється черезрядкова розгортка у прогресивну?*

## СЛОВНИК ТЕРМІНІВ І СКОРОЧЕНЬ

**АДІКМ** – адаптивна диференціальна імпульсно-кодова модуляція (англ. ADPCM) – різновид ДІКМ, у якій крок квантування й параметри передбачення змінюються залежно від поточних властивостей кодованого сигналу.

**АЦП** – аналого-цифровий перетворювач.

**БіХ-фільтр** – фільтр із нескінченною імпульсною характеристикою (рекурсивний цифровий фільтр).

**Буферний ЗП** – запам'ятовувальний пристрій для тимчасового зберігання даних. Зокрема, у кодерах і декодерах MPEG використовується для узгодження цих пристроїв з каналом передавання кодованих відеоданих з постійною швидкістю.

**Вектор руху** (вектор зсуву) – пара чисел, що виражають розраховані в результаті оцінки руху зсуву блоку зображення за двома просторовими координатами.

**Відеопослідовність** (англ. *video sequence*) – найбільша структурна одиниця потоку відеоданих у стандартах MPEG-1, MPEG-2. Зустрічається також назва «відеоряд».

**Відеостик** – назва інтерфейсів для передавання цифрових телевізійних сигналів відповідно до Рекомендації ITU-R BT 656, у якій визначено паралельний і послідовний відеостик.

**Відповідність блоків** (англ. *block matching*) – один з методів оцінювання руху, за яким для кожного блока (макроблока) кодованого зображення знаходять в опорному зображенні ділянку таких же розмірів, яка найбільше відповідає за обраним критерієм цьому блоку.

**Вейвлет-перетворення** – ортогональне одновимірне або двовимірне перетворення, у результаті якого вхідний сигнал

розділяється на складові з низькою роздільною здатністю і з високою роздільною здатністю.

**Гамма-корекція** – нелінійне перетворення телевізійних сигналів, яке виконується з метою корекції нелінійності передавальних характеристик різних вузлів телевізійної системи, наприклад кінескопа. Також гамма-корекція сприяє зменшенню впливу квантування на якість зображення.

**Гібридне кодування** – метод кодування рухомих зображень, що застосовується у стандартах MPEG-1, MPEG-2 і MPEG-4. Поєднує кодування всередині кадру та міжкадрове кодування.

**Деквантування** – операція, обернена квантуванню, у результаті якої збільшується кількість рівнів квантування, якими подаються значення відліків цифрового сигналу або іншої інформації.

**Декодування** – операція, обернена кодуванню, у результаті якої інформація набуває вигляду, який вона мала до відповідного кодування.

**Декомпресія** – операція, обернена компресії.

**Демодуляція** – операція, обернена модуляції, у результаті якої відновлюється модульований сигнал.

**Децимація** – див. проріджування.

**ДІКМ** – диференціальна імпульсно-кодова модуляція (англ. DPCM), також називають «кодування з передбаченням» – вид ІКМ, у якій на основі значень одного або декількох попередніх відліків сигналу формується передбачене значення поточного відліку, а каналом зв'язку передається похибка передбачення – різниця дійсного й передбаченого значень поточного відліку сигналу.

**Дискретизація** – подання неперервного сигналу у вигляді послідовності його значень (відліків або вибірок), що надходять через певні інтервали часу. У процесі дискретизації зображення подається

у вигляді матриці відліків, які певним чином розміщені у площині зображення.

**Дискретна частота** – частота дискретного сигналу, вимірювана в частках частоти дискретизації.

**ДКП** – дискретне косинусне перетворення.

**ДПФ** – дискретне перетворення Фур'є.

**Елемент зображення** – див. піксел.

**Ентропійне кодування** – кодування інформації, що дає змогу зменшити її обсяг без необоротних втрат, ґрунтується на використанні статистичних властивостей джерела повідомлення, наприклад на розподілі ймовірностей появи різних символів, слів тощо.

**ІКМ** – імпульсно-кодова модуляція (англ. PCM), – спосіб передавання інформації в цифровій формі, тобто цифровий код кожного відліку сигналу передається каналом зв'язку у вигляді послідовності імпульсів.

**Інтерактивне телебачення** – телевізійні системи, у яких глядач може впливати на інформацію, яку він одержує, передаючи сигнали зворотному каналом у зворотному напрямку на головну станцію.

**Інтерполяція** 1. Перетворення дискретного сигналу в неперервний, за допомогою заповнення за певним правилом проміжків часу (або у разі зображень – проміжків простору) між відліками дискретного сигналу. 2. Збільшення кількості відліків дискретного сигналу через введення між відліками вихідного дискретного сигналу додаткових відліків, значення яких визначаються за заданими правилами.

**Кадр** 1. У телебаченні кадрами називають передані й відтворені одне за одним нерухомі зображення. Якщо частота передавання кадрів досить велика, глядач сприймає зображення як таке, що

неперервно рухається. Під час передавання кадри розкладаються на рядки. 2. У багатьох випадках кадрами називають структурні одиниці потоку даних, наприклад, звукові кадри.

**Кадрове кодування** – один з режимів кодування відеоінформації з черезрядковою розгорткою у стандартах MPEG-1 і MPEG-2, у якому кожний макроблок містить елементи зображення обох полів кадру.

**Канальне кодування** – кодування, яке виконується безпосередньо перед передаванням інформації каналом зв'язку та зазвичай поєднується з модуляцією з метою узгодження з параметрами каналу зв'язку, підвищення перешкодостійкості тощо.

**Квадратурна амплітудна маніпуляція (КАМн або КАМ)** – один з видів модуляції, який застосовується під час передавання цифрових сигналів, за якої дискретно змінюються амплітуди двох квадратурних складових ( $\cos$  і  $\sin$ ) несучої.

**Квантування** 1. Перетворення неперервного або дискретного сигналу за допомогою заміни кожного його значення найближчим рівнем квантування. 2. Під час квантування сигналу, вже поданого у цифровій формі, може зменшуватися кількість бітів на кожний відлік сигналу.

**КІХ-фільтр** – фільтр з кінцевою імпульсною характеристикою (нерекурсивний цифровий фільтр).

**Кодек** – пристрій, що може виконувати функції як кодера, так і декодера.

**Кодер** – пристрій, у якому відбувається кодування.

**Кодування** 1. Подання результату квантування кожного відліку дискретизованого сигналу двійковим числом за певним правилом. 2. Перетворення інформації завдяки заміні груп символів на інші групи символів за певними правилами. Метою кодування можуть бути

зменшення надмірності або стиснення інформації, підвищення перешкодостійкості, захист інформації від доступу небажаних осіб тощо.

**Кодування всередині кадру** (англ. *intraframe*) – стиснення відеоінформації, що ґрунтується на зменшенні кадрової надмірності.

**Кодування за Хаффманом** – варіант кодування зі змінною довжиною кодових слів, що забезпечує найменшу надмірність передаваної інформації.

**Кодування зі змінною довжиною кодових слів** – один з варіантів ентропійного кодування, за якого зменшення обсягу передаваної інформації досягається за рахунок того, що більш імовірні символи подаються більш короткими кодовими словами.

**Кодування з передбаченням** – див. ДІКМ.

**Кольорорізницеві сигнали** – сигнали в кольоровому телебаченні, що містять інформацію про колір зображення та не впливають на його яскравість.

**Компенсація руху** – формування зображення з блоків, елементів, фрагментів іншого (опорного) зображення, які зміщено на вектори руху, що визначені в результаті оцінки руху. Використовується для формування передбачених кадрів у стандартах MPEG.

**Композитне кодування** – перетворення повного кольорового телевізійного сигналу (ПКТС) у цифрову форму.

**Компонентне кодування** – роздільне перетворення сигналу яскравості та кольорорізницевих сигналів у цифрову форму.

**Компресія** 1. Стиснення інформації за допомогою ефективного кодування, вилучення несуттєвої інформації тощо. 2. Стиснення динамічного діапазону сигналу за допомогою нелінійного перетворення.

**Лінійне передбачення** (англ. *Linear Prediction*) – формування

передбаченого значення сигналу (див. ДІКМ) у вигляді лінійної комбінації декількох попередніх значень.

**Макроблок** – прямокутна ділянка зображення розміром 16×16 пікселів, що є основною структурною одиницею кодування зображень у стандартах MPEG. Для макроблоків виконуються оцінка й компенсація руху.

**Масштабованість** (англ. *Scalability*) – властивість методів кодування й синтаксису потоку даних MPEG-2, що дає змогу отримувати зображення з неповною якістю з частини потоку відеоданих. У стандарті MPEG-4 масштабованість поширено і на звукове супровідження.

**МДКП** – модифіковане дискретне косинусне перетворення, різновид ДКП, яке використовується в MPEG-1, MPEG-2.

**Медіанна фільтрація** – фільтрація, за якою відлік сигналу на виході визначають як медіану розподілу значень декількох відліків вхідного сигналу. Медіанна фільтрація може бути як одновимірною, так і двовимірною, тобто просторовою.

**Міжкадрове кодування** (англ. *interframe*) – стиснення відеоінформації, яке полягає у використанні міжкадрової кореляції.

**Модем** – пристрій, у якому виконуються модуляція та демодуляція.

**Модуляція** – змінювання одного або декількох параметрів електричного сигналу, який називають несучим коливанням або просто несучою, відповідно до значень модулювального сигналу.

**Надмірність зображення** – наявність у зображенні складових, які можуть бути усунути без істотного погіршення якості зображення, що сприймається візуально.

**Нерекурсивний цифровий фільтр** – цифровий фільтр, у якому значення сформованого відліку сигналу на виході залежить лише від



значень відліків вхідного сигналу.

**Об'єкт** – основний елемент об'єктно-орієнтованого підходу. У стандарті MPEG-4 використано поняття «відеооб'єкт», яким може бути фрагмент зображення або ціле зображення; та «аудіовізуальний об'єкт», у якому поєднуються відеооб'єкт і аудіооб'єкт.

**Опорне зображення** (англ. *Reference Picture*) – кадр або поле, за яким виконується передбачення з компенсацією руху макроблоків кодованого зображення.

**Оцінка руху** – визначення зсувів окремих блоків, елементів, фрагментів зображення відносно їх положень в іншому (опорному) зображенні. За результатами оцінки руху може виконуватися компенсація руху.

**Похибка передбачення** – різниця між передбаченим й дійсним значеннями сигналу. Під час кодування зображень похибкою передбачення фрагмента (макроблока) зображення може бути матриця чисел, кожний елемент якої дорівнює різниці значень сигналів відповідних елементів передбаченого і дійсного фрагментів.

**Піксел** (піксель) (англ. *pixel*) – елемент дискретного зображення, у межах якого яскравість і колірність є постійними.

**Поле** – половина кадру під час черезрядкової розгортки, що містить непарні рядки кадру (перше поле) або парні рядки кадру (друге поле). Поле передається й відтворюється за один період вертикальної розгортки.

**Польове кодування** – один з режимів кодування відеоінформації з черезрядковою розгорткою у стандартах MPEG-1 і MPEG-2, за якого макроблок складається з елементів зображення лише першого або лише другого поля.

**Перетворення Хаара** – один з різновидів вейвлет-перетворення.

**Прогресивна розгортка** – варіант розгортки, за якої всі рядки

передаються послідовно один за одним.

**Пропущений макроблок** (англ. *skipped*) – макроблок, що не передається, тому що збігається з відповідним макроблоком опорного зображення.

**Проріджування** – зменшення кількості відліків цифрового сигналу, у результаті якого залишають, наприклад, кожний 2-й, або кожний 3-й, кожний 4-й і т. д. відлік, а інші відкидають.

**Просторова фільтрація** – перетворення зображення, за якого яскравість і колір кожного елемента зображення на виході визначаються як певна функція від яскравості та кольору елементів початкового зображення.

**Просторова частота** (для неперервних зображень) – величина, обернена просторовому періоду, яка показує, скільки періодів зміни якого-небудь параметра, наприклад яскравості, укладається на одиниці довжини. Одиниця вимірювання –  $\text{мм}^{-1}$ .

**Просторовий (двовимірний) фільтр** – пристрій, за допомогою якого виконується просторова фільтрація.

**Профіль** – в стандартах MPEG-2 і MPEG-4 визначає складність методів кодування й наявність масштабованості.

**ПКТС** – повний кольоровий телевізійний сигнал – сигнал аналогового кольорового телебачення, що містить сигнал яскравості, колірну несучу, модульовану кольорорізними сигналами, синхроімпульси, гасильні імпульси та інші компоненти.

**Рекомендація 601** – повна назва – Рекомендація ITU-R BT 601. Стара назва – Рекомендація 601 МККР або CCIR-601. Визначає параметри цифрового подання телевізійних сигналів телебачення звичайної чіткості.

**Рекомендація ITU-R BT-709-3** – документ, що визначає параметри цифрового подання сигналів телебачення високої чіткості

(ТВЧ), загальні для Європи, США та Японії.

**Рекурсивний цифровий фільтр** – цифровий фільтр, у якому значення сформованого відліку сигналу на виході залежить як від значень відліків вхідного сигналу, так і від значень раніше сформованих відліків сигналу на виході.

**Рівень** (англ. *Level*) – у стандарті MPEG-2 визначено рівні за роздільною здатністю зображення.

**Сигнал яскравості** – сигнал, що містить інформацію про яскравість усіх елементів зображення та відповідає відеосигналу чорно-білого телебачення.

**Синхроімпульси** – імпульси, що входять до складу повного телевізійного сигналу з метою синхронізації розгортки у телевізійному приймачі з відповідними розгортками передавальної телевізійної камери.

**Скремблер** – пристрій, у якому виконується скремблювання.

**Скремблювання** – додавання сигналу, що заважає, до передаваного сигналу для захисту інформації від несанкціонованого доступу.

**Слайс** (англ. *slice*) – група макроблоків, які надходять один за одним у процесі кодування. Для всіх макроблоків слайса задається однаковий параметр стиснення, що записується у заголовок слайса.

**Сцена** – одне з понять стандарту MPEG-4, що являє собою сукупність аудіовізуальних об'єктів.

**ТВЧ** – телебачення високої чіткості.

**Текстура** – дрібна структура зображення, заповнення контурів об'єктів.

**Фазова маніпуляція (ФМн)** – один з видів модуляції, який застосовується під час передавання цифрових сигналів та полягає в дискретній зміні фази несучої.

**ЦАП** – цифро-аналоговий перетворювач.

**Цифровий фільтр** – пристрій, що перетворює послідовність відліків вхідного цифрового сигналу в послідовність відліків цифрового сигналу на виході.

**ЦПОС** (англ. DSP) – цифровий процесор обробки сигналів.

**Частота дискретизації** – величина, обернена періоду надходження відліків дискретного сигналу. Відповідно до теореми Котельникова, частота дискретизації має щонайменше у 2 рази перевищувати верхню граничну частоту дискретизованого сигналу.

**Черезрядкова розгортка** – варіант розгортки, за якої спочатку передаються всі непарні рядки кадру, які утворюють перше поле, а потім – усі парні рядки кадру, що утворюють друге поле. Черезрядкова розгортка застосовується у системах телевізійного мовлення.

**Швидкість передавання двійкових символів** (англ. *Bitrate*) – кількість бітів інформації, яка передається за секунду. Одиниці вимірювання – біт/с, кбіт/с, Мбіт/с тощо.

**ШПФ** – швидке перетворення Фур'є – алгоритми прискореного виконання ДПФ.

**Шум квантування** – похибка, що виникає в результаті квантування сигналу.

**ADPCM** – Adaptive Differential Pulce Code Modulation – див. АДІКМ.

**BER** – Bit Error Rate – частота помилок на 1 біт переданої інформації.

**B-frame** (*B*-кадр) (від Bidirectional – двоспрямований) – кадр, макроблоки якого можуть кодуватися з передбаченням за двома опорними зображеннями, одне з яких розташовано перед *B*-кадром, а інше – після нього.

**CIF** – Common Intermediate Format – формат кадру, що містить 288 рядків по 352 елемента в рядку для сигналу яскравості й 176 рядків по 144 елемента в рядку для кожного з кольорорізницевих сигналів.

**Content-based coding** – кодування, що спирається на зміст. Один з варіантів кодування зображень у стандарті MPEG-4, у якому можуть кодуватися об'єкти складної форми.

**DCT** – див. ДКП.

**DMIF** – Delivery Multimedia Integration Framework (інтегрована система доставки мультимедіа) – протокол, що забезпечує керування потоками даних у стандарті MPEG-4.

**DPCM** – Differential Pulse Code Modulation – див. ДІКМ.

**DVB** – Digital Video Broadcasting (цифрове відеомовлення) – система цифрового телебачення, що розвивається в Європі, включає кабельне (DVB-C), супутникове (DVB-S), наземне (DVB-T), а в перспективі й інші види телевізійного мовлення.

**Field** – див. поле.

**Frame** – див. кадр.

**GOP** – Group of Pictures – група зображень. У стандартах MPEG-1 і MPEG-2 група зображень (кадрів або полів), які надходять одне за одним, що починається з *I*-кадру.

**H.261, H.262, H.263** – Рекомендації ITU-T, які описують методи кодування відеоінформації для систем відеозв'язку.

**HDTV** – High-Definition Television – див. ТВЧ.

**I-frame** (*I*-кадр) (від англ. *Intraframe* – всередині кадру) – кадр, з якого починається GOP у стандартах MPEG-1 і MPEG-2. Всі макроблоки *I*-кадру кодуються всередині кадру.

**ISO** – International Organization for Standardization – Міжнародна

організація зі стандартизації, заснована у 1947 році, включає більше 100 національних організацій зі стандартизації.

**ITU** – International Telecommunication Union – Міжнародний союз електрозв'язку (МСЕ). Одна з функцій цієї організації – розроблення Рекомендацій, які фактично є міжнародними стандартами.

**ITU-R** – підрозділ ITU, що відповідає за розроблення рекомендацій у галузях радіозв'язку й телебачення.

**ITU-T** – підрозділ ITU, що відповідає за розроблення рекомендацій у галузях телефонії та цифрового передавання даних.

**JPEG** – Joint Photographic Experts Group – об'єднана група експертів по фотографії, робоча група ISO, що займається розробленням стандартів кодування та стиснення нерухомих зображень, які носять таку ж назву.

**Level** – див. рівень.

**Linear Prediction** – див. лінійне передбачення.

**Lossless JPEG** – JPEG без утрат. Метод стиснення зображень, розроблений групою JPEG, ґрунтується на кодуванні з передбаченням, дає можливість отримати істотно менше стиснення, ніж «звичайний» JPEG.

**LPC** – Linear Predictive Coding – кодування методом лінійного передбачення (див. ДІКМ, лінійне передбачення).

**Motion JPEG** – метод стиснення рухомих зображень, за якого кожний кадр стискається методом JPEG незалежно від інших кадрів, використовують його передусім у системах відеомонтажу.

**MPEG** – Moving Picture Expert Group – група експертів із рухомих зображень. Робоча група ISO, що займається розробленням стандартів кодування і стиснення відео- та аудіоданих. Назва групи наявна у назвах стандартів.

**MPEG-1** – стандарт стиснення рухомих зображень і звукового

супроводження, затверджений у 1993 році. Його здебільшого використовують під час записування відеопрограм формату SIF на лазерні диски, забезпечуючи стиснення до 1,5 Мбіт/с.

**MPEG-2** – стандарт стиснення рухомих зображень, і звукового супроводження, затверджений у 1994 році, який є основою сучасних систем цифрового телебачення.

**MPEG-3** – проект групи MPEG щодо стиснення зображень ТВЧ, що увійшов у MPEG-2 і нині окремо не існує.

**MPEG-4** – стандарт стиснення рухомих зображень і звукового супроводження. Прийнятий у 1998–1999 роках, у першу чергу спрямований на передавання аудіовізуальної інформації каналами зв'язку з вузькою частотною смугою пропускання та інтерактивні телевізійні системи.

**MPEG-7** – проект групи MPEG, основною метою якого є стандартизація засобів опису змісту відео- та аудіоінформації.

**OFDM** – Orthogonal Frequency Division Multiplex (ортогональне частотне мультиплексування) – один з видів модуляції, яку застосовують у системах цифрового телебачення.

**PCM** – Pulse Code Modulation – див. ІКМ.

**P-frame** (*P*-кадр) (від англ. *Predictive* – передбачений) – кадр, макроблоки якого можуть кодуватися з передбаченням за опорним зображенням, у якості якого може використовуватися попередній *I*-кадр або *P*-кадр.

**PIP** – Picture in picture – зображення в зображенні (кадр у кадрі).

**Pixel** – Picture element – елемент зображення (див. піксел).

**Profile** – див. профіль.

**QAM** – див. квадратурна амплітудна маніпуляція.

**QCIF** – Quarter Common Intermediate Format – формат кадру, що

становить 1/4 відносно формату CIF, тобто 176×144 елементів сигналу яскравості.

**Scalability** – див. масштабованість.

**Skipped macroblock** – див. пропущений макроблок.

**SIF** 1. Standard Interchange Format – формат кадру, що містить 288 рядків по 352 елементи зображення в рядку або 240 рядків по 352 елементи зображення в рядку. 2. Source Input Format – формат кадру, що містить 240 рядків по 320 елементів зображення в рядку.

**SQCIF** – Sub-Quarter Common Interchange Format – формат кадру, що містить 128×96 елементів сигналу яскравості.

**VLBV** – Very Low Bitrate Video – рівень дуже низької швидкості передавання даних у стандарті MPEG-4.

**VLC** – Variable Length Coding – див. кодування зі змінною довжиною кодових слів.

**Wavelet-Transform** – див. вейвлет-перетворення.

**4:2:2; 4:2:0; 4:1:1; 4:4:4** – варіанти форматів дискретизації сигналу яскравості та кольорорізнiceвих сигналів.



## СПИСОК ЛІТЕРАТУРИ

1. Телевидение [Текст]: учеб. для студ. высш. учеб. заведений / В. Е. Джакония, А. А. Гоголь, Я. В. Друзин и др.; под ред. В. Е. Джакония. – 2-е изд. – М.: Горячая линия-Телеком, 2002. – 640 с. – Библиогр.: с. 633–635. – 2000 экз. – ISBN 5-93517-070-1.
2. Быков Р. Е. Теоретические основы телевидения [Текст]: учеб. для студ. высш. учебн. заведений / Р. Е. Быков. – СПб.: Лань, 1998. – 288 с.
3. Смирнов А. В. Основы цифрового телевидения [Текст]: учеб. пособие для студ. высш. учеб. заведений / А. В. Смирнов. – М.: Горячая линия-Телеком, 2001. – 224 с. – Библиогр.: с. 221–222. – 3000 экз. – ISBN 5-93517-059-0.
4. Смирнов А. В. Цифровое телевидение: от теории к практике [Текст] / А. В. Смирнов, А. Е. Пескин. – М.: Горячая линия-Телеком, 2005. – 352 с. – Библиогр.: с. 347–349. – 3000 экз. – ISBN 5-93517-222-4.
5. Грязин Г. Н. Системы прикладного телевидения [Текст]: учеб. пособие для студ. высш. учебн. завед. / Г. Н. Грязин. – СПб.: Политехника, 2000. – 277 с. – Библиогр.: с. 273–275. – 1000 экз. – ISBN 5-7325-0445-1.
6. Цифровое преобразование изображений [Текст]: учебн. пособие для студ. высш. учеб. заведений / Р. Е. Быков, Р. Фрайер, К. В. Иванов, А. А. Манцветов; под ред. Р. Е. Быкова. – М.: Горячая линия-Телеком, 2003. – 228 с. – Библиогр.: с. 221–222. – 3000 экз. – ISBN 5-93517-119-8.
7. Куприянов М. С. Цифровая обработка сигналов: процессоры, алгоритмы, средства проектирования [Текст] / М. С. Куприянов, Б. Д. Матюшкин. – 2-е изд. – СПб.: Политехника, 1999. – 592 с.

8. ISO/IEC DIS 10918-1. Information Technology – Digital Compression and Coding of Continuous-tone Still Images: Requirements and Guidelines./Ed/1, JTS 1/SC 9, 1994.
9. ISO/IEC 13818-1. Information Technology – Coding of Moving Pictures and Associated Audio Information. Part 1: Systems./ Ed/1, JTS 1/SC29, 1994.
10. ISO/IEC 13818-2. Information Technology – Coding of Moving Pictures and Associated Audio Information. Part 2: Video./ Ed/1, JTS 1/SC 29, 1994.
11. ISO/IEC 13818-3. Information Technology – Coding of Moving Pictures and Associated Audio Information. Part 3: Audio./ Ed/1, JTS 1/SC 29, 1994.
12. ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio. MPEG-4. Overview. 1999.
13. ITU-T Recommendation H.261. Line Transmission of Non-Telephone Signals. Video Codec for Audiovisual Services at px64 kbits. 03/93.
14. ITU-T Recommendation H.263. Transmission of Non-Telephone Signals. Video Coding for Low Bit Rate Communication. 03/96.
15. ITU-T Recommendation H.262. Transmission of Non-Telephone Signals. Information Technology - Generic Coding Of Moving Pictures and Associated Audio Information: Video. 07/95.
16. Яковлев А. Н. Введение в вейвлет-преобразования [Текст]: учеб. пособие для студ. высш. учеб. заведений / А. Н. Яковлев. – Новосибирск: НГТУ, 2003. – 104 с. – Библиогр.: с. 102–103. – 500 экз. – ISBN 5-7782-0361-6.
17. Мартынов Н.Н. Введение в MATLAB 6.x [Текст] / Н. Н. Мартынов. – М.: КУДИЦ ОБРАЗ, 2002. – 352 с. – Библиогр.: с. 348. – 3000 экз. – ISBN 5-93378-039-1.